Save and Load Machine Learning Models in Python with scikit-learn

by **Jason Brownlee** on June 8, 2016 in **Python Machine Learning**

Last Updated on February 4, 2020

Finding an accurate machine learning model is not the end of the project.

In this post you will discover how to save and load your machine learning model in Python using scikit-learn.

This allows you to save your model to file and load it later in order to make predictions.

Discover how to prepare data with pandas, fit and evaluate models with scikit-learn, and more in my new book, with 16 step-by-step tutorials, 3 projects, and full python code.

Let's get started.

> **Update Jan/2017**: Updated to reflect changes to the scikit-learn API in version 0.18.
> **Update Mar/2018**: Added alternate link to download the dataset as the original appears to have been taken down.
> **Update Oct/2019**: Fixed typo in comment.
> **Update Feb/2020**: Updated joblib API.

Save and Load Machine Learning Models in Python with scikit-learn
Photo by Christine, some rights reserved.

Finalize Your Model with pickle

Pickle is the standard way of serializing objects in Python.

You can use the pickle operation to serialize your machine learning algorithms and save the serialized format to a file.

Later you can load this file to deserialize your model and use it to make new predictions.

The example below demonstrates how you can train a logistic regression model on the Pima Indians onset of diabetes dataset, save the model to file and load it to make predictions on the unseen test set (update: download from here).

```
# Save Model Using Pickle
import pandas
from sklearn import model_selection
from sklearn.linear_model import LogisticRegression
import pickle
url = "https://raw.githubusercontent.com/jbrownlee/Datasets/master/pima-indians-diabetes.
names = ['preg', 'plas', 'pres', 'skin', 'test', 'mass', 'pedi', 'age', 'class']
dataframe = pandas.read_csv(url, names=names)
array = dataframe.values
X = array[:,0:8]
Y = array[:,8]
test_size = 0.33
seed = 7
X_train, X_test, Y_train, Y_test = model_selection.train_test_split(X, Y, test_size=test_
# Fit the model on training set
model = LogisticRegression()
model.fit(X_train, Y_train)
```

```
# save the model to disk
filename = 'finalized_model.sav'
pickle.dump(model, open(filename, 'wb'))

# some time later...

# load the model from disk
loaded_model = pickle.load(open(filename, 'rb'))
result = loaded_model.score(X_test, Y_test)
print(result)
```

Running the example saves the model to **finalized_model.sav** in your local working directory. Load the saved model and evaluating it provides an estimate of accuracy of the model on unseen data.

```
0.755905511811
```

Finalize Your Model with joblib

Joblib is part of the SciPy ecosystem and provides utilities for pipelining Python jobs.

It provides utilities for saving and loading Python objects that make use of NumPy data structures, efficiently.

This can be useful for some machine learning algorithms that require a lot of parameters or store the entire dataset (like K-Nearest Neighbors).

The example below demonstrates how you can train a logistic regression model on the Pima Indians onset of diabetes dataset, saves the model to file using joblib and load it to make predictions on the unseen test set.

```
# Save Model Using joblib
import pandas
from sklearn import model_selection
from sklearn.linear_model import LogisticRegression
import joblib
url = "https://raw.githubusercontent.com/jbrownlee/Datasets/master/pima-indians-diabetes.
names = ['preg', 'plas', 'pres', 'skin', 'test', 'mass', 'pedi', 'age', 'class']
dataframe = pandas.read_csv(url, names=names)
array = dataframe.values
X = array[:,0:8]
Y = array[:,8]
test_size = 0.33
seed = 7
X_train, X_test, Y_train, Y_test = model_selection.train_test_split(X, Y, test_size=test_
# Fit the model on training set
model = LogisticRegression()
model.fit(X_train, Y_train)
# save the model to disk
filename = 'finalized_model.sav'
joblib.dump(model, filename)

# some time later...

# load the model from disk
loaded_model = joblib.load(filename)
result = loaded_model.score(X_test, Y_test)
print(result)
```

Running the example saves the model to file as **finalized_model.sav** and also creates one file for each NumPy array in the model (four additional files). After the model is loaded an estimate of the accuracy of the model on unseen data is reported.

```
0.755905511811
```

Tips for Finalizing Your Model

This section lists some important considerations when finalizing your machine learning models.

Python Version. Take note of the python version. You almost certainly require the same major (and maybe minor) version of Python used to serialize the model when you later load it and deserialize it.
Library Versions. The version of all major libraries used in your machine learning project almost certainly need to be the same when deserializing a saved model. This is not limited to the version of NumPy and the version of scikit-learn.
Manual Serialization. You might like to manually output the parameters of your learned model so that you can use them directly in scikit-learn or another platform in the future. Often the algorithms used by machine learning algorithms to make predictions are a lot simpler than those used to learn the parameters can may be easy to implement in custom code that you have control over.

Take note of the version so that you can re-create the environment if for some reason you cannot reload your model on another machine or another platform at a later time.

Summary

In this post you discovered how to persist your machine learning algorithms in Python with scikit-learn.

You learned two techniques that you can use:

The pickle API for serializing standard Python objects.
The joblib API for efficiently serializing Python objects with NumPy arrays.

Do you have any questions about saving and loading your machine learning algorithms or about this post? Ask your questions in the comments and I will do my best to answer them.

About Jason Brownlee

Jason Brownlee, PhD is a machine learning specialist who teaches developers how to get results with modern machine learning methods via hands-on tutorials.

View all posts by Jason Brownlee →

« Binary Classification Tutorial with the Keras Deep Learning Library

Regression Tutorial with the Keras Deep Learning Library in Python »

254 Responses to *Save and Load Machine Learning Models in Python with scikit-learn*

Kayode October 18, 2016 at 6:15 pm #

REPLY

Thank you so much for this educative post.

Jason Brownlee October 19, 2016 at 9:17 am #

REPLY

You're welcome Kayode.

Ehsan Arabnezhad May 12, 2020 at 5:07 pm #

REPLY

How can I predict in a case when there are difference between model's and test data's columns? I mean in case of one-hot encoding step that has been done before?

Jason Brownlee May 13, 2020 at 6:28 am #

REPLY

If a label has not been seen before, you can ignore it, e.g. encode it as all zeros.

Urjit January 8, 2020 at 12:26 am #

REPLY

Hey, i trained the model for digit recognition but when i try to save the model i get the following error. Please help.

import pickle

# save the model to disk
filename = 'digit_model.sav'
pickle.dump(model, open(filename, 'wb'))
#saved_model=pickle.dumps(model)

ERROR-

TypeError Traceback (most recent call last)
in
3 # save the model to disk
4 filename = 'digit_model.sav'
—-> 5 pickle.dump(model, open(filename, 'wb'))
6 #saved_model=pickle.dumps(model)

TypeError: can't pickle _thread._local objects

Jason Brownlee January 8, 2020 at 8:27 am #

REPLY

Sorry to hear that, perhaps try posting your code and error on stackoverflow?

Williams Abodunrin March 27, 2020 at 9:34 pm #

REPLY

Is it a must the finalised model is saved with a ".sav" file extension. Can we save it as a python file(.py)

Jason Brownlee March 28, 2020 at 6:17 am #

REPLY

No, it is binary, not Python code.

TonyD November 13, 2016 at 3:52 pm #

REPLY

Hi Jason,

I have two of your books and they are awesome. I took several machine learning courses before, however as you mentioned they are more geared towards theory than practicing. I devoured your Machine Learnign with Python book and 20x my skills compared to the courses I took.

I found this page by Googling a code snippet in chapter 17 in your book. The line:

loaded_model = pickle.load(open(filename, ‘rb’))

throws the error:

runfile(‘C:/Users/Tony/Documents/MassData_Regression_Pickle.py’,
wdir=’C:/Users/Tony/Documents’)
File “C:/Users/Tony/Documents/MassData_Regression_Pickle.py”, line 55
loaded_model = pickle.load(open(filename, ‘rb’))
^
SyntaxError: invalid syntax

Jason Brownlee November 14, 2016 at 7:36 am #

REPLY

Thanks TonyD.

I wonder if there is a copy-paste error, like an extra space or something?

Does the code example (.py file) provided with the book for that chapter work for you?

William January 7, 2019 at 9:37 pm #

REPLY

As Jason already said, this is a copy paste problem. In your line specifically, the quotes are the problem.

```
loaded_model = pickle.load(open(filename, ‘rb’))
```

It should be

```
loaded_model = pickle.load(open(filename, 'rb'))
```

Try to understand the difference :).

Jason Brownlee January 8, 2019 at 6:49 am #

REPLY

Thanks.

This might help:

https://machinelearningmastery.com/faq/single-faq/how-do-i-copy-code-from-a-tutorial

REPLY

John January 8, 2020 at 1:15 pm #

Hey TonyD
Can you please share the books with me if you don't mind.

I'm very eager to learn machine learning but i can't afford to buy the books. If you could help me out with the books it would be great.

REPLY

Jason Brownlee January 8, 2020 at 2:26 pm #

You can discover my best free tutorials here:
https://machinelearningmastery.com/start-here/

REPLY

Konstantin November 19, 2016 at 6:01 am #

Hello, Jason

Where we can get X_test, Y_test "sometime later"? It is "garbag collected"!
X_test, Y_test not pickled In your example you pickle classifier only but you keep refer to x and y. Real applications is not single flow I found work around and get Y from clf.classes_ object.

What is correct solution? Should we pickle decorator class with X and Y or use pickled classifier to pull Ys values? I didn't find legal information from documentation on KNeighborclassifier(my example) as well; how to pull Y values from classifier.

Can you advise?

REPLY

Jason Brownlee November 19, 2016 at 8:51 am #

Hi Konstantin,

I would not suggest saving the data. The idea is to show how to load the model and use it on new data – I use existing data just for demonstration purposes.

You can load new data from file in the future when you load your model and use that new data to make a prediction.

If you have the expected values also (y), you can compare the predictions to the expected values and see how well the model performed.

REPLY

Guangping Zhang November 21, 2016 at 6:01 am #

I'm newer Pythoner, your code works perfect! But where is the saved file? I used

windows 10.

REPLY

Jason Brownlee November 22, 2016 at 6:56 am #

Thanks Guangping.

The save file is in your current working directory, when running from the commandline.

If you're using a notebook or IDE, I don't know where the file is placed.

REPLY

Mohammed Alnemari December 13, 2016 at 2:45 pm #

Hi Jason ,
I am just wondering if can we use Yaml or Json with sklearn library . I tried to do it many times but I could not reach to an answer . I tried to do it as your lesson of Kares , but for some reason is not working . hopefully you can help me if it is possible

REPLY

Jason Brownlee December 14, 2016 at 8:24 am #

Hi Mohammed, I believe the serialization of models to yaml and json is specific to the Keras library.

sklearn serialization is focused on binary files like pickle.

REPLY

Normando Zubia December 29, 2016 at 9:55 am #

Hi, my name is Normando Zubia and I have been reading a lot of your material for my school lessons.

I'm currently working on a model to predict user behavoir in a production environment. Due to several situations I can not save the model in a pickle file. Do you know any way to save the model in a json file?

I have been playing a little with sklearn classes and I noticed that if I save some parameters for example: n_values_, feature_indices_ and active_features_ in a OneHotEncoding model I can reproduce the results. Could this be done with a pipeline? Or do you think I need to save each model's parameters to load each model?

PS: Sorry for my bad english and thanks for your attention.

REPLY

Jason Brownlee December 30, 2016 at 5:49 am #

Hi Normando,

If you are using a simple model, you could save the coefficients directly to file. You can then try

and put them back in a new model later or implement the prediction part of the algorithm yourself (very easy for most methods).

Let me know how you go.

Samuel February 6, 2017 at 3:14 pm #

REPLY

Hello Jason,

I am new to machine learning. I am your big fan and read a lot of your blog and books. Thank you very much for teaching us machine learning.

I tried to pickle my model but fail. My model is using VGG16 and replace the top layer for my classification solution. I further narrowed down the problem and find that it is the VGG16 model failed to pickle. Please find my simplified code below and error log below:

It will be highly appreciated if you can give me some direction on how to fix this error.

Thank you very much

———————————————————-

```
# Save Model Using Pickle
from keras.applications.vgg16 import VGG16
import pickle

model = VGG16(weights='imagenet', include_top=False)

filename = 'finalized_model.sav'
pickle.dump(model, open(filename, 'wb'))
```

———————————————————-

```
/Library/Frameworks/Python.framework/Versions/2.7/bin/python2.7
/Users/samueltin/Projects/bitbucket/share-card-ml/pickle_test.py
Using TensorFlow backend.
Traceback (most recent call last):
File "/Users/samueltin/Projects/bitbucket/share-card-ml/pickle_test.py", line 8, in
pickle.dump(model, open(filename, 'wb'))
File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 1376, in dump
Pickler(file, protocol).dump(obj)
File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 224, in dump
self.save(obj)
File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 331, in save
self.save_reduce(obj=obj, *rv)
File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 425, in
save_reduce
save(state)
File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 286, in save
f(self, obj) # Call unbound method with explicit self
File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 655, in
save_dict
self._batch_setitems(obj.iteritems())
File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 669, in
_batch_setitems
save(v)
```

```
  File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 286, in save
  f(self, obj) # Call unbound method with explicit self
  File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 606, in
save_list
  self._batch_appends(iter(obj))
  File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 621, in
_batch_appends
  save(x)
  File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 331, in save
  self.save_reduce(obj=obj, *rv)
  File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 425, in
save_reduce
  save(state)
  File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 286, in save
  f(self, obj) # Call unbound method with explicit self
  File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 655, in
save_dict
  self._batch_setitems(obj.iteritems())
  File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 669, in
_batch_setitems
  save(v)
  File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 286, in save
  f(self, obj) # Call unbound method with explicit self
  File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 606, in
save_list
  self._batch_appends(iter(obj))
  File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 621, in
_batch_appends
  save(x)
  File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 331, in save
  self.save_reduce(obj=obj, *rv)
  File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 425, in
save_reduce
  save(state)
  File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 286, in save
  f(self, obj) # Call unbound method with explicit self
  File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 655, in
save_dict
  self._batch_setitems(obj.iteritems())
  File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 669, in
_batch_setitems
  save(v)
  File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 331, in save
  self.save_reduce(obj=obj, *rv)
  File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 425, in
save_reduce
  save(state)
  File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 286, in save
  f(self, obj) # Call unbound method with explicit self
  File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 655, in
```

File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 655, in save_dict
self._batch_setitems(obj.iteritems())
File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 669, in _batch_setitems
save(v)
File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 331, in save
self.save_reduce(obj=obj, *rv)
File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 425, in save_reduce
save(state)
File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 286, in save
f(self, obj) # Call unbound method with explicit self
File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 655, in save_dict
self._batch_setitems(obj.iteritems())
File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 669, in _batch_setitems
save(v)
File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 286, in save
f(self, obj) # Call unbound method with explicit self
File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 655, in save_dict
self._batch_setitems(obj.iteritems())
File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 669, in _batch_setitems
save(v)
File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 286, in save
f(self, obj) # Call unbound method with explicit self
File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 606, in save_list
self._batch_appends(iter(obj))
File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 621, in _batch_appends
save(x)
File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 331, in save
self.save_reduce(obj=obj, *rv)
File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 425, in save_reduce
save(state)
File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 286, in save
f(self, obj) # Call unbound method with explicit self
File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 655, in save_dict
self._batch_setitems(obj.iteritems())
File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 669, in _batch_setitems
save(v)
File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 331, in save
self.save_reduce(obj=obj, *rv)
File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 425, in

```
 File "/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py", line 425, in
save_reduce
 save(state)
 File “/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py”, line 286, in save
 f(self, obj) # Call unbound method with explicit self
 File “/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py”, line 655, in
save_dict
 self._batch_setitems(obj.iteritems())
 File “/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py”, line 669, in
_batch_setitems
 save(v)
 File “/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py”, line 331, in save
 self.save_reduce(obj=obj, *rv)
 File “/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py”, line 425, in
save_reduce
 save(state)
 File “/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py”, line 286, in save
 f(self, obj) # Call unbound method with explicit self
 File “/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py”, line 655, in
save_dict
 self._batch_setitems(obj.iteritems())
 File “/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py”, line 669, in
_batch_setitems
 save(v)
 File “/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py”, line 286, in save
 f(self, obj) # Call unbound method with explicit self
 File “/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py”, line 606, in
save_list
 self._batch_appends(iter(obj))
 File “/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py”, line 621, in
_batch_appends
 save(x)
 File “/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py”, line 286, in save
 f(self, obj) # Call unbound method with explicit self
 File “/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py”, line 568, in
save_tuple
 save(element)
 File “/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py”, line 286, in save
 f(self, obj) # Call unbound method with explicit self
 File “/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py”, line 655, in
save_dict
 self._batch_setitems(obj.iteritems())
 File “/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py”, line 669, in
_batch_setitems
 save(v)
 File “/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/pickle.py”, line 306, in save
 rv = reduce(self.proto)
 File “/Library/Frameworks/Python.framework/Versions/2.7/lib/python2.7/copy_reg.py”, line 70, in
_reduce_ex
 raise TypeError, “can’t pickle %s objects” % base.__name__
```

TypeError: can't pickle module objects

Process finished with exit code 1

REPLY

Jason Brownlee February 7, 2017 at 10:11 am #

Sorry Samuel, I have not tried to save a pre-trained model before. I don't have good advice for you.

Let me know how you go.

REPLY

huikang September 21, 2018 at 11:50 am #

Is there a more efficient method in machine learning than joblib.load(), storing the model directly in memory and using it again?

REPLY

Jason Brownlee September 21, 2018 at 2:21 pm #

Sure, you can make an in-memory copy. I think sklearn has a clone() function that you can use.

REPLY

Amy March 8, 2017 at 7:03 am #

I have trained a model using liblinearutils. The model could not be saved using pickle as it gives error that ctype module with pointers cannot be pickled. How can I save my model?

REPLY

Jason Brownlee March 8, 2017 at 9:47 am #

Sorry Amy, I don't have any specific examples to help.

Perhaps you can save the coefficients of your model to file?

REPLY

SHUBHAM BHARDWAJ April 3, 2017 at 10:42 pm #

Thanks a lot, very useful

REPLY

Jason Brownlee April 4, 2017 at 9:14 am #

You're welcome!

Benju April 11, 2017 at 1:35 am #

REPLY

My saved modells are 500MB+ Big….is that normal?

Jason Brownlee April 11, 2017 at 9:34 am #

REPLY

Ouch, that does sound big.

If your model is large (lots of layers and neurons) then this may make sense.

Anupam April 13, 2017 at 2:32 am #

REPLY

How to use model file (“finalized_model.sav”) to test unknown data. Like, if the model is for tagger , how this model will tag the text file data? Is there any example?

Jason Brownlee April 13, 2017 at 10:08 am #

REPLY

You can load the saved model and start making predictions (e.g. yhat = model.predict(X)).

See this post on finalizing models:
http://machinelearningmastery.com/train-final-machine-learning-model/

Oss Mps April 21, 2017 at 3:09 pm #

REPLY

Dear Sir, please advice on how to extract weights from pickle dump? Thank you

Jason Brownlee April 22, 2017 at 9:23 am #

REPLY

I would suggest extracting coefficients from your model directly and saving them in your preferred format.

Suhas May 24, 2017 at 4:44 am #

REPLY

Hi I love your website; it’s very useful!

Are there any examples showing how to save out the training of a model after say 100 epochs/iterations? It’s not immediately clear from looking at joblib or scikit learn.

This is esp. useful when dealing with large datasets and/or computers or clusters which may be

unreliable (e.g., subject to system reboots, etc.)

REPLY

Jason Brownlee May 24, 2017 at 4:59 am #

I'm not sure how to do this with sklearn. You may need to write something custom. Consider posting to stackoverflow.

REPLY

Viktor May 30, 2017 at 8:52 am #

Hey!
Is it possible to open my saved model and make a prediction on cloud server where is no sklearn installed?

REPLY

Jason Brownlee June 2, 2017 at 12:31 pm #

no.

You could save the coefficients from within the model instead and write your own custom prediction code.

REPLY

Clemence June 8, 2017 at 6:55 pm #

Hello Jason and thank you very much, it's been very helpful.

Do you know if it's possible to load features transformation with the ML model?
I'm mostly thinking of categorical variables that we need to encode into numerical ones.

I'm using sklearn to do that, but I don't know if we can (as for Spark), integrate this transformation with the ML model into the serialized file (Pickle or Joblib).

```
#Encode categorical variable into numerical ones
from sklearn.preprocessing import LabelEncoder
list_var = ['country', 'city']

encoder = LabelEncoder()
for i in list_var:
    df[i] = encoder.fit_transform(df[i])
```

Then I fit the model on the training dataset…

And I need to save this transformation with the model. Do you know if that's possible ?
Thank you!

REPLY

Jason Brownlee June 9, 2017 at 6:23 am #

I'm not sure I follow, sorry

I'm not sure I follow sorry.

You can transform your data for your model, and you can apply this same transform in the future when you load your model.

You can save the transform objects using pickle. Is that what you mean?

Bhavani Shanker June 22, 2017 at 1:24 am #

REPLY

Hi Jason,
Kindly accept my encomiums for the illustrative lecture that you have delivered on Machine Learning using Python.

****

```
# save the model to disk
filename = 'finalized_model.sav'
joblib.dump(model, filename)

# sometime later…

# load the model from disk
loaded_model = joblib.load(filename)
result = loaded_model.score(X_test, Y_test)
print(result)
```

****

After saving the model 'finalized_model.sav' , How can recall the saved model in the new session at later date?

I would appreciate if you can advice on this

Jason Brownlee June 22, 2017 at 6:11 am #

REPLY

The code after "sometime later" would be in a new session.

jinsh June 28, 2017 at 8:57 pm #

REPLY

Hello sir,

The above code saves the model and later we can check the accuracy also
but what i have to do for making predicting the class of unknown data?
I mean which function have to be called ?

eg: 2,132,40,35,168,43.1,2.288,33

can you suggest how to get the class of above data through prediction ?

thank you

Jason Brownlee June 29, 2017 at 6:35 am #

REPLY

Pass in input data to the predict function and use the result.

```
yhat = model.predict(X)
```

Ukesh Chawal July 24, 2017 at 11:09 pm #

REPLY

Can we use “pickling” to save an LSTM model and to load or used a hard-coded pre-fit model to generate forecasts based on data passed in to initialize the model?

When I tried to use it, it gave me following error:

PicklingError: Can’t pickle : attribute lookup module on builtins failed

Jason Brownlee July 25, 2017 at 9:44 am #

REPLY

No.

See this tutorial on how to save Keras models:
http://machinelearningmastery.com/save-load-keras-deep-learning-models/

Ukesh Chawal July 25, 2017 at 11:59 pm #

REPLY

Great. It worked.

You are awesome Jason. Appreciated.

Jason Brownlee July 26, 2017 at 7:57 am #

REPLY

Glad to hear it.

akatsuki August 9, 2017 at 1:21 pm #

REPLY

tbh this is best of the sites on web. Great!
I love the email subscriptions of yours as a beginner they are quite helpful to me .

Jason Brownlee August 10, 2017 at 6:45 am #

REPLY

Thanks, I’m glad to hear that.

vikash August 10, 2017 at 9:32 pm #

REPLY

Hi @Jason Brownlee thanks for such informative blog. Can you please guide me for a problem where i would like to retrain the .pkl model only with new dataset with new class keeping the previous learning intact. I had thought that model.fit(dataset,label) will do that but it forgets the previous learning. Please suggest me some techniques for it.
Thanks

Jason Brownlee August 11, 2017 at 6:42 am #

REPLY

Sorry, I don't follow. Can you please restate your question?

sassashi August 28, 2017 at 4:41 am #

REPLY

Hi Jason, I believe @vikash is looking for a way to continuously train the model with new examples after the initial training stage. This is something I am searching for as well. I know it is possible to retrain a model in tensorflow with new examples but I am not sure if it's possible with sklearn.

to expand the question some more: 1-you train a model with sklearn 2-save it with pickle or joblib
3-then you get your hands on some new examples that were not available at the time of initial training "step 1" 4-you load the previous model 5-and now you try to train the model again using the new data without losing the previous knowledge… is step 5 possible with sklearn?

Jason Brownlee August 28, 2017 at 6:52 am #

REPLY

I have not updated a model in sklearn, but I would expect you can.

Here is an example of updating a model in Keras which may help in general principle:
https://machinelearningmastery.com/update-lstm-networks-training-time-series-forecasting/

Navdeep Singh August 22, 2017 at 8:30 pm #

REPLY

Hi Json,

I need your guidance on Updation of saved pickle files with new data coming in for training

I recall 3 methods, Online Learning which is train one every new observation coming in and in this case model would always be biased towards new features ,which i dont wana do

Second is, Whenever some set of n observations comes, embedd it with previous data and do retraining again from scratch, that i dont want to do as in live environment it will take lot of time

Third is Mini batch learning, i know some of algorithm like SGD and other use partial fit method and

Third is Mini batch learning, I know some of algorithm like SGD and other use partial fit method and do same but I have other algorithms as week like random forest , decision tress, logistic regression. I wana ask can i update the previously trained pickle with new training ?

I am doing it in text classification, I read that possibly doing this, model update pickle will not take new features of new data (made using tfidf or countvectorizer) and it would be of less help.

Also as domain is same, and If client(Project we are working for) is different , inspite of sharing old data with new client (new project), could i use old client trained model pickle and update it with training in new client data. Basically I am transferring learning

REPLY

Jason Brownlee August 23, 2017 at 6:48 am #

Great question.

This is a challenging problem to solve. Really, the solution must be specific to your project requirements.

A flexible approach may be to build-in capacity into your encodings to allow for new words in the future.

The simplest approach is to ignore new words.

These, and other strategies are testable. See how performance degrades under both schemes with out-of-band test data.

REPLY

Merari September 11, 2017 at 7:59 am #

Gracias por compartir,
Existe alguna forma en la que pueda realizar predicciones con nuevos datos solo con el modelo guardado? llamando este modelo desde un archivo nuevo? lo he intentado con la instruccion final:

```
# load the model from disk
 loaded_model = pickle.load(open(filename, 'rb'))
 result = loaded_model.score(X_test, Y_test)
 print(result)
```

pero no lo he logrado

373/5000
Thanks for sharing,
Is there any way I can make predictions with new data only with the saved model? calling this model from a new file? I have tried with the final instruction:

```
# load the model from disk
 loaded_model = pickle.load (open (filename, 'rb'))
 result = loaded_model.score (X_test, Y_test)
 print (result)
```

but I have not achieved it

Jason Brownlee September 11, 2017 at 12:11 pm #

REPLY

That is exactly what we do in this tutorial.

What is the problem exactly?

AP September 29, 2017 at 6:36 am #

REPLY

Hi Jason, I learn a lot reading your python books and blogs. Thank you for everything.

I'm having an issue when I work on text data with loaded model in a different session. I fit and transform training data with countvectorizer and tfidf. Then I only transform the test data with the fitted instances as usual. But, when work on loaded pretrained model in a different session, I am having problem in feature extraction. I can't just transform the test data as it asks for fitted instance which is not present in the current session. If I fit and transform on test data only, model prediction performance drastically decreases. I believe that is wrong way of doing machine learning. So, how can I do the feature extraction using countvectorizer, tfidf or other cases while working with previously trained model?

I'm using spark ML but I think it would be the same for scikit-learn as well.

Jason Brownlee September 30, 2017 at 7:31 am #

REPLY

Perhaps you can pickle your data transform objects as well, and re-use them in the second session?

Bhavya Chugh October 29, 2017 at 5:57 am #

REPLY

Hi Jason,

I trained a random forest model and saved the same as a pickle file in my local desktop. I then copied that pickle file to my remote and tested the model with the same file and it is giving incorrect predictions. I am using python 3.6 in my local and python 3.4 in my remote, however the version of scikit-learn are same. Any ideas why this may be happening?

Jason Brownlee October 29, 2017 at 6:00 am #

REPLY

No idea, perhaps see if the experiment can be replicated on the same machine? or different machines with the same version of Python?

Berkin Albert Antony November 10, 2017 at 5:45 pm #

REPLY

Hi Jason Brownlee,

I have a LogisticRegression model for binary classification. I wish to find a similar data points in a trained model for a given test data points. So that I can show these are similar data points predicted with these same class.

Could you please suggest your thoughts for the same. I am using scikit learn logistic regression

Thanks

Jason Brownlee November 11, 2017 at 9:18 am #

REPLY

Perhaps you could find data points with a low Euclidean distance from each other?

James November 16, 2017 at 8:47 am #

REPLY

Hi Jason –

If you pickle a model trained on a subset of features, is it possible to view these features after loading the pickled model in a different file? For example: original df has features a,b,c,d,e,f. You train the model on a,c,e. Is it possible to load the pickled model in a separate script and see the model was trained on a,c,e?

Thanks,
James

Jason Brownlee November 16, 2017 at 10:33 am #

REPLY

Yes, you can save your model, load your model, then use it to make predictions on new data.

Mrinal Mitra November 22, 2017 at 6:26 am #

REPLY

Hi Jason,

Thanks for explaining it so nicely. I am new to this and will be needing your guidance. I have data using which I have trained the model. Now I want this model to predict an untested data set. However, my requirement is an output which will have the data and corresponding prediction by the model. For example, record 1 – type a, record 2 – type a, record 3 – type c and so on. Could you please guide me on this?

Jason Brownlee November 22, 2017 at 11:16 am #

REPLY

You can provide predictions one at a time or in a group to the model and the predictions will be in the same order as the inputs.

Does that help?

Lais February 21, 2020 at 2:24 am #

REPLY

Hi Jason!
I am new in python and I have the same problem. I already have trained the model to predict sex, but I need to predict an untested dataset (input a .csv file with the same features).
Please, what command should I have to use?

Cam February 21, 2020 at 2:58 am #

REPLY

My untested dataset has 14 features, does it makes a sense to do predictions one at a time? I tried the other tutorial (https://machinelearningmastery.com/how-to-connect-model-input-data-with-predictions-for-machine-learning/), but it still have problems:

in _validate_X_predict
X = check_array(X, dtype=DTYPE, accept_sparse="csr")

ValueError: at least one array or dtype is required

Sorry if it is a silly question (I've been looking for the sequence of commands to predict new data for hours). Thanks for your help!

Jason Brownlee February 21, 2020 at 8:27 am #

See this:
https://machinelearningmastery.com/make-predictions-scikit-learn/

Jason Brownlee February 21, 2020 at 8:26 am #

REPLY

See this for making predictions:
https://machinelearningmastery.com/make-predictions-scikit-learn/

Niranjan December 3, 2017 at 3:22 pm #

REPLY

Hi,

I am using chunks functionality in the read csv method in pandas and trying to build the model iteratively and save it. But it always saves the model that is being built in the last chunk and not the entire model. Can you help me with it

clf_SGD = SGDClassifier(loss='modified_huber', penalty='l2', alpha=1e-3, max_iter=500, random_state=42)

```
pd.read_csv("file_name",chunksize = 1000):
"""
data preparation and cleaning
"""
hashing = hv.fit_transform(X_train['description'])
    clf_SGD.partial_fit(hashing, y_train, classes= y_classes)

joblib.dump(clf_SGD, source_folder + os.path.sep+'text_clf_sgd.pkl')
```

Jason Brownlee December 4, 2017 at 7:46 am #

REPLY

Sorry, I'm not sure I follow, could you please try reframing your question?

Shabbir December 8, 2017 at 8:50 am #

REPLY

Hi Jason,
This is extremely helpful and saved me quite a bit of processing time.

I was training a Random Forest Classifier on a 250MB data which took 40 min to train everytime but results were accurate as required. The joblib method created a 4GB model file but the time was cut down to 7 Minutes to load. That was helpful but the results got inaccurate or atleast varied quite a bit from the original results. I use average of 2 Decision Tree and 1 Random Forest for the model. Decision Tree Models have kept there consistency loading vs training but RF hasn't. Any ideas?

Nilanka December 19, 2017 at 9:10 pm #

REPLY

Thank you very useful!!

Jason Brownlee December 20, 2017 at 5:43 am #

REPLY

You're welcome.

Gokhan December 28, 2017 at 2:55 pm #

REPLY

Hello, if i load model
loaded_model = joblib.load(filename)
result = loaded_model.score(X_test, Y_test)
print(result)

can i use this model for another testsets to prediction?

REPLY

Jason Brownlee December 29, 2017 at 5:17 am #

Sure.

REPLY

Vinay Boddula January 20, 2018 at 5:31 am #

Hi Jason,

How do I generated new X_Test for prediction ? This new X_Test needs to make sure that the passed parameters are same in the model was trained with.

Background: I am basically saving the model and predicting with new values from time to time. How do we check whether the new values have all the parameters and correct data type.

REPLY

Jason Brownlee January 20, 2018 at 8:25 am #

Visualization and statistics.

I have many posts on the topic, try the search box.

REPLY

Sekar February 1, 2018 at 4:06 am #

Jason. Very good article. As asked by others, in my case I am using DecisionTreeClassifier with text feature to int transformation. Eventhough, you mentioned that transformation map can also be picked and read back, is there any example available? Will it be stored in the same file or it will be another file?

REPLY

Jason Brownlee February 1, 2018 at 7:24 am #

In a separate file.

REPLY

Yousif February 5, 2018 at 8:01 pm #

Thank you so much professor
we get more new knowledge

REPLY

Jason Brownlee February 6, 2018 at 9:12 am #

You're welcome. Also, I'm not a professor.

REPLY

Adarsh C February 8, 2018 at 12:29 pm #

REPLY

HI sir,

I would like to save predicted output as a CSV file. After doing ML variable I would like to save “y_predicted”. And I’m using python ide 3.5.x I have pandas,sklearn,tensorflow libraries

Jason Brownlee February 9, 2018 at 8:58 am #

REPLY

You can save the numpy array as a csv.
https://docs.scipy.org/doc/numpy-1.13.0/reference/generated/numpy.savetxt.html

Atul March 11, 2018 at 6:45 am #

REPLY

Hi Jason,

I would like to save predicted output as a CSV file. After doing ML variable I would like to save “y_predicted”. How I can save Naive Bayes, SVM, RF and DT Classification for final predictions for all samples saved as a .csv with three columns namely Sample, Actual value, Prediction values

Jason Brownlee March 12, 2018 at 6:24 am #

REPLY

Perhaps create a dataframe with all the columns you require and save the dataframe directly via to_csv():
https://pandas.pydata.org/pandas-docs/stable/generated/pandas.DataFrame.to_csv.html

Tommy March 22, 2018 at 11:14 pm #

REPLY

I have a list of regression coefficients from a paper. Is there a way to load these coefficients into the sklearn logistic regression function to try and reproduce their model?
Thanks!
Tommy

Jason Brownlee March 23, 2018 at 6:07 am #

REPLY

No model is needed, use each coefficient to weight the inputs on the data, the weighted sum is the prediction.

Vincent April 10, 2018 at 10:25 am #

REPLY

Hi,all

I am using scikit 0.19.1
I generated a training model using random forest and saved the model. These were done on ubuntu 16.01 x86_64.
I copied the model to a windows 10 64 bit machine and wanted to reuse the saved model. But unfortunately i get the following
Traceback (most recent call last):
File “C:\Users\PC\Documents\Vincent\nicholas\feverwizard.py.py”, line 19, in
rfmodel=joblib.load(modelfile)
File “C:\Python27\lib\site-packages\sklearn\externals\joblib\numpy_pickle.py”, line 578, in load
obj = _unpickle(fobj, filename, mmap_mode)
File “C:\Python27\lib\site-packages\sklearn\externals\joblib\numpy_pickle.py”, line 508, in _unpickle
obj = unpickler.load()
File “C:\Python27\lib\pickle.py”, line 864, in load
dispatchkey
File “C:\Python27\lib\pickle.py”, line 1139, in load_reduce
value = func(*args)
File “sklearn\tree_tree.pyx”, line 601, in sklearn.tree._tree.Tree.cinit
ValueError: Buffer dtype mismatch, expected ‘SIZE_t’ but got ‘long long’

What could be happening? Is it because of a switch from ubuntu to windows? However i am able to reuse the model in my ubuntu.

REPLY

Jason Brownlee April 11, 2018 at 6:29 am #

Perhaps the pickle file is not portable across platforms?

REPLY

Pramod April 17, 2018 at 9:03 pm #

Can we load model trained on 64 bit system on 32 bit operating system..?

REPLY

Jason Brownlee April 18, 2018 at 8:04 am #

I’m skeptical that it would work. Try it and see. Let me know how you go.

REPLY

Arnaud April 17, 2018 at 9:29 pm #

Dear Jason :

Thank you for ‘le cours’ which is very comprehensive.

I have a maybe tricky but ‘could be very usefull’ question about my newly created standard Python object.

Is it possible to integrate a call to my Python object in a Fortran program ?

Basically I have a deterministic model in which I would like to make recursive calls to my Python

object at every time step.

Do I need some specific libraries ?

Thank you
Best regards

REPLY

Jason Brownlee April 18, 2018 at 8:06 am #

You're welcome.

I suspect it is possible. It's all just code at the end of the day. You might need some kind of Python-FORTRAN bridge software. I have not done this, sorry.

REPLY

Pratip April 23, 2018 at 4:32 pm #

Hi Sir ,
I wanted to know if its possible to combine the scikit preloaded datasets with some new datasets to get more training data to get further higher accuracy or firstly run on the scikit loaded dataset and then save model using pickle an run it on another dataset .
Which method will be correct ?
Please help .

REPLY

Jason Brownlee April 24, 2018 at 6:20 am #

Sure, you can, but it may only make sense if the data was collected in the same way from the same domain.

REPLY

swati kumari April 12, 2019 at 5:07 pm #

How it can be done? When I am loading the pickle and try to fit new data , the model gets fitted with new data only.

REPLY

Jason Brownlee April 13, 2019 at 6:24 am #

If the model has already been fit, saved, loaded and is then trained on new data, then it is being updated, not trained from scratch.

Perhaps I don't understand the problem you're having?

REPLY

Ishit Gandhi May 4, 2018 at 6:00 pm #

Hii Jason,

Can you put example of how to store and load Pipeline models?

eg.

clf = Pipeline([(“rbm”,rbm),(“logistic”,logistic)])
clf.fit(trainX,trainY)

REPLY

Jason Brownlee May 5, 2018 at 6:18 am #

Perhaps use pickle? This might help:
https://machinelearningmastery.com/save-load-machine-learning-models-python-scikit-learn/

REPLY

Akash May 14, 2018 at 4:15 pm #

Hi jason,
My name is Akash Joshi.I am trying to train my scikit svm model with 101000 images but I run out of memory.Is there a way where I can train the svm model in small batches?Can we use pickle?

REPLY

Jason Brownlee May 15, 2018 at 7:51 am #

Perhaps try running on a machine with more RAM, such as an EC2 instance?

Perhaps try using a sample of your dataset instead?

Perhaps use a generator to progressively load the data?

REPLY

Samarth May 14, 2018 at 4:54 pm #

Hi Jason

I want to know how can presist a minmax transjformation? There are ways to persist the final model but to persist the transformations?

Thanks

REPLY

Jason Brownlee May 15, 2018 at 7:51 am #

Save the min and max values for each variable.

Or save the whole object.

REPLY

SOORAJ T S May 16, 2018 at 12:30 am #

thank you the post, it is very informative but i have a doubt about the labels or names of the dataset can specify each.

REPLY

Jason Brownlee May 16, 2018 at 6:05 am #

What do you mean exactly?

REPLY

SOORAJ T S May 16, 2018 at 4:11 pm #

names = [‘preg’, ‘plas’, ‘pres’, ‘skin’, ‘test’, ‘mass’, ‘pedi’, ‘age’, ‘class’]

in the above code what are these “preg” , “plas”, “pres” etc…

REPLY

Jason Brownlee May 17, 2018 at 6:24 am #

You can learn about these features here:
https://github.com/jbrownlee/Datasets/blob/master/pima-indians-diabetes.names

REPLY

SOORAJ T S May 17, 2018 at 4:23 pm #

thank you sir…

REPLY

Aniko June 7, 2018 at 12:13 am #

HI Jason!

I created a machine learning (GBM) model to predict house prices and a Django application to usability. This model has more than 1000 n_estimators and it takes more than 1 minutes to load before getting the prediction in every request.
I would like to load joblib dump file just once and store the model in memory, avoiding loading the model in every get requests.

What is your best practice for this?

Thanks

REPLY

Jason Brownlee June 7, 2018 at 6:31 am #

This sounds like a web application software engineering question rather than a machine learning question.

Perhaps you can host the model behind a web service?

REPLY

Aniko June 7, 2018 at 6:51 pm #

thank you, meanwhile I found some caches -related solution in Django documentation, this perhaps solve the loading problem

REPLY

Jason Brownlee June 8, 2018 at 6:07 am #

Glad to hear it.

REPLY

LamaOS223 June 9, 2018 at 2:00 pm #

okay what if i had 2 datasets for Example a Loan datasets
the first dataset has a Loan_Status attribute
and the second one does not have a Loan_Status attribute
if i trained the model on the first dataset and i want to predict the Loan_Status for the second dataset, how to do that? please make it simple for me i'm beginner

REPLY

Jason Brownlee June 10, 2018 at 5:58 am #

This process wills how you how to work through a predictive model systematically:
https://machinelearningmastery.com/start-here/#process

REPLY

Imti July 12, 2018 at 4:55 pm #

Hey Jason, I am working on a model to classify text files. I am using the CountVectorizer, TfidfTransformer and SGDClassifier in the same sequence on a set of data files. I am saving the SGDClassifier object via the joblib.dump method you have mentioned in this article.

Do I also need to save the vectorizer and transformer objects/models ? Since when i take a new file for classification I will need to go through these steps again.

REPLY

Jason Brownlee July 13, 2018 at 7:33 am #

Yes, they are needed to prepare any data prior to using the model.

REPLY

Dennis Faucher July 28, 2018 at 2:38 am #

Just what I needed today. Thank syou.

REPLY

Jason Brownlee July 28, 2018 at 6:38 am #

I'm happy to hear that Dennis.

REPLY

Tejaswini July 30, 2018 at 9:01 am #

Hi Jason,
Appreciate for the article. when i am saving the model and loading it in different page.Then it is showing different accuracy.

Problem trying to solve: I am using oneclasssvm model and detecting outliers in sentences.

REPLY

Jason Brownlee July 30, 2018 at 2:15 pm #

I have not seen that, are you sure you are evaluating the model on exactly the same data?

REPLY

Tejaswini August 2, 2018 at 2:10 pm #

Yes Jason i am using gensim word2vec to convert text into feature vectors and then performing classification task.after saving model and reloading in another session its giving different results.

REPLY

Jason Brownlee August 2, 2018 at 2:11 pm #

That is odd. I have not seen this.

Perhaps report a fault/bug?

REPLY

EvapStudent August 7, 2018 at 1:36 am #

Hi Jason,

I am training a neural network using MLPRegressor, trying to predict pressure drop in different geometries of heat exchangers. I think I have gotten the network to train well with low MRE, but I can't figure out how to use the network. When I tried to load using pickle and call again, I am getting an error when using "score". I am new to python so not sure how to go about bringing in new data for the network to predict or how to generalize doing so.

Jason Brownlee August 7, 2018 at 6:28 am #

REPLY

I don't recommend using pickle. I recommend using the Keras API to save/load your model.

Once you find a config that works for your problem, perhaps switch from the sklearn wrappers to the Keras API directly.

EvapStudent August 7, 2018 at 11:13 pm #

REPLY

Hi Jason,

Thanks for the recommendation. Is there no easy way to save a model and call from it to use in scikit learn? I have been getting good results with the model I have made on there, I just don't know how to get it to the point where I can actually use the network (i.e. put in a geometry and get it's predictions).

If using Keras API to save/load is the best option, how do I go about doing that?

Jason Brownlee August 8, 2018 at 6:21 am #

REPLY

There may be, but I don't have an example, sorry.

Golnoush August 21, 2018 at 1:38 am #

REPLY

Hello Jason,

Thank you for your nice tutorial! Does pickle.dump(model, open(filename, 'wb')) only save the neural network model or it also save the parameters and weights of the model?
Does the back propagation and training is done again when we use pickle.load ?
What I would like to do is that I aim to save the whole model and weights and parameters during training and use the same trained model for every testing data I have. I would be so thankful if you could assist me in this way.

Jason Brownlee August 21, 2018 at 6:19 am #

REPLY

I believe you cannot use pickle for neural network models – e.g. Keras models.

Somo August 29, 2018 at 3:05 pm #

REPLY

Hi Jason,

I am trying to save my model using joblib.dump(model, 'model.pkl') and load it back up in another .py

file model = joblib.load('model.pkl') but then the accuracy dropped and each time I run it the accuracy differs a lot. I coefficient and the intercept and the same for both models. Any ideas why this might happen. Thanks in advance.

REPLY

Jason Brownlee August 30, 2018 at 6:26 am #

Perhaps this will help:
https://machinelearningmastery.com/faq/single-faq/why-do-i-get-different-results-each-time-i-run-the-code

REPLY

Dhrumil September 1, 2018 at 3:11 pm #

Hey man I am facing a trouble with pickle, when I try to load my .pkl model I am getting following error :

UnicodeDecodeError: 'ascii' codec can't decode byte 0xbe in position 3: ordinal not in range(128)

Can you please tell me something since I have tried all fixes I could find..

REPLY

Jason Brownlee September 2, 2018 at 5:30 am #

Perhaps post your error on stackoverflow?

REPLY

Aakash Aggarwal September 8, 2018 at 4:57 am #

I want to develop to train my model and save in pickle file. From the next time onwards, when i want to train the model, it should save in previously created pickle file in append mode that reduces the time of training the model. I am using LogisticRegression model.

Any helps would be greatly appreciated.

REPLY

Jason Brownlee September 8, 2018 at 6:17 am #

This post shows how:
https://machinelearningmastery.com/save-load-machine-learning-models-python-scikit-learn/

REPLY

Aakash Aggarwal October 2, 2018 at 12:45 am #

This article shows how to save a model that is built from scratch. But I am looking to train the model by including additional data so as to achieve high prediction performance and accuracy for unseen data. Is there any leads or approach you can think?

Jason Brownlee October 2, 2018 at 6:26 am #

REPLY

I don't understand, sorry. Training a model and saving it are separate tasks.

My3 October 15, 2018 at 10:11 pm #

REPLY

Hi Jason,

I have some requirement to integrate python code with Java.

I have a ML model which is trained as saved as pickle file, Randomforestclassifier.pkl. I want to load this one time using java and then execute my "prediction" part code which is written python. So my workflow is like:

1. Read Randomforestclassifier.pkl file (one time)
2. Send this model as input to function defined in "python_file.py" which is executed from java for each request
3. python_file.py has prediction code and predictions returned should be captured by java code

Please provide suggestions for this workflow requirement I have used processbuilder in java to execute python_file.py and everything works fine except for model loading as one time activity.

Can you help me with some client server python programming without using rest APIs for one time model loading?

Jason Brownlee October 16, 2018 at 6:37 am #

REPLY

I recommend treating it like any other engineering project, gather requirements, review options, minimize risk.

Rahul October 18, 2018 at 6:10 pm #

REPLY

Hi Jason,My3,

I have a similar requirement to integrate java with python as my model is in python and in my project we are using java.

Could you please help here.

Jason Brownlee October 19, 2018 at 6:01 am #

REPLY

Thanks for the suggestion.

Theekshana October 30, 2018 at 12:35 am #

REPLY

Hi Jason,

I have trained my model and evaluated the accuracy using cross-validation score.

After evaluating the model, should I train my model with the whole data set and then save the new trained model for new future data. (assuming the new model performs with good accuracy around mean accuracy from cross-validation)

Thank you for your tutorials and instant replies to questions.

Jason Brownlee October 30, 2018 at 6:03 am #

REPLY

Yes, see this post:
https://machinelearningmastery.com/train-final-machine-learning-model/

Gagan December 11, 2018 at 5:56 pm #

REPLY

Jason, thanks so much for value add.

Jason Brownlee December 12, 2018 at 5:50 am #

REPLY

You're welcome.

Roger January 1, 2019 at 5:55 am #

REPLY

That helped me a lot. Thank you

Jason Brownlee January 1, 2019 at 6:29 am #

REPLY

I'm happy to hear that.

Kiril Kirov January 4, 2019 at 3:19 am #

REPLY

How would you go about saving and loading a scikit-learn pipeline that uses a custom function created using FunctionTransformer?

Jason Brownlee January 4, 2019 at 6:33 am #

REPLY

Perhaps pickle?

Shubham January 4, 2019 at 8:37 am #

REPLY

Hey Jason,

I have a very basic question, let's say I have one model trained on 2017-2018, and then after 6 months I feel to retrain it on new data. What does retraining actually means here, do I need to have target for my new data and needs to trained from scratch for new time period, I obviously don't have the target and then how model will learn from new data.

Jason Brownlee January 4, 2019 at 11:01 am #

REPLY

You have many options, e.g. develop a new model, update the old model, some mixture of the two with an ensemble.

Rajesh Mahajan January 18, 2019 at 7:44 am #

REPLY

Hi Jason,

I am new to this.. So pardon, if I am asking something incorrect…

I have two stages. Build model and predict.

For Build model:
I am using vectorizer.fit_transform(data) and building the logistic model. My data is a bunch of comments and the target is a set of categories. In order for me to use that model for predicting categories for new comments, I am using the vector created earlier during building of model to predict

So, when I do the save model joblib.dump(log_model, "model.sav")

Foe Predict:
When I try to re-run the model (saved) at a later point of time, I don't have the original vectorizer anymore with the original data set

```
log_model = joblib.load("model.sav")
inputfeatures_nd = vectorizer.transform(newComment);
pred = log_model.predict(inputfeatures_nd)
```

I get this error – sklearn.exceptions.NotFittedError: CountVectorizer – Vocabulary wasn't fitted.

What do you suggest I should do ? Should I be serializing the vector also and storing ?

Jason Brownlee January 18, 2019 at 10:15 am #

REPLY

You must use the same vectorizer that was used when training the model. Save it along with your model.

Rajesh Mahajan January 18, 2019 at 1:17 pm #

REPLY

Thanks Jason! Yes it worked after I save and reload.

Jason Brownlee January 19, 2019 at 5:32 am #

REPLY

I'm happy to hear that.

Ahmed Sahlol February 13, 2019 at 8:23 pm #

REPLY

Thanks Jason for your intersting subjects.
I have this error when saving a VGG16 model after training and testing on my own dataset (can't pickle _thread.RLock objects) when applying the two methods. I also read somewhere that Keras models are not Pickable. So, do you think that those methods are applicable in my case?

Jason Brownlee February 14, 2019 at 8:43 am #

REPLY

Yes, don't use pickle for Keras models.

Use the Keras save API:
https://machinelearningmastery.com/save-load-keras-deep-learning-models/

shashank February 15, 2019 at 12:24 am #

REPLY

Awesome post! Keep doing the good work bro!

Jason Brownlee February 15, 2019 at 8:06 am #

REPLY

Thanks.

Nick February 19, 2019 at 2:31 am #

REPLY

Hi Jason, your promised "free" book never came, it looks you are collecting emails for promotions

Jason Brownlee February 19, 2019 at 7:27 am #

REPLY

Sorry to hear that, I can confirm that your email is not in the system, perhaps a typo when you entered it?

Nevertheless, email me directly and I will send you whichever free ebook you are referring to:
https://machinelearningmastery.com/contact/

mayank March 6, 2019 at 7:31 am #

REPLY

Hi Jason, I have a .sav file where my random forest model has been trained. I have to get back the whole python script for training the model from that .sav file. Is it possible??

Jason Brownlee March 6, 2019 at 8:02 am #

REPLY

You can load the saved model and start using it.

Rimsha March 29, 2019 at 5:05 am #

REPLY

Hi Jason, I have trained a model of Naved Baise for sentiment analysis through a trained dataset file of .csv and now I want to use that model for check sentiments of the sentences which are also saved in another .csv file, how could I use?

Jason Brownlee March 29, 2019 at 8:44 am #

REPLY

Save the model, then load it in a new example and make predictions.

Nisha March 29, 2019 at 11:23 pm #

REPLY

Hi Jason,

I have a Class Layer defined to do some functions in Keras. I trained the model and pickled it. Now when I try to unpickle it, I see an error saying- unknown layer Layer.
How should be pickle the model in this case?

Jason Brownlee March 30, 2019 at 6:28 am #

REPLY

I don't recommend using pickle for Keras models, instead Keras has it's own save model functions:
https://machinelearningmastery.com/save-load-keras-deep-learning-models/

Sevval May 8, 2019 at 5:56 pm #

REPLY

Can i use my previously saved model for prediction ?

Jason Brownlee May 9, 2019 at 6:38 am #

REPLY

Yes, load it and call model.predict().

See this post:
https://machinelearningmastery.com/how-to-make-classification-and-regression-predictions-for-deep-learning-models-in-keras/

Sara May 13, 2019 at 1:37 pm #

REPLY

When in this article you say:
"You might manually output the parameters of your learned model so that you can use them directly in scikit-learn",
I see that we can manually get the tuned hyperparameters, or for example in svm, we can get weight coefficients (coef_),
BUT, is it possible to get svm hyperplane parameters, w and b (y=wx+b) for future predictions?
I believe that pickle saves and load the learned model including w and b but is there any way we can manually output w and b and see what they are in scikit learn?
Many thanks

Jason Brownlee May 13, 2019 at 2:33 pm #

REPLY

I believe they will be accessible as attributes within the SVM class.

You might need to take a closer look at the API or even the source code to dig out the coefficients and how they are specifically used by sklearn to make a prediction.

Thank god for open source though, it's all there for us!

Samuel May 18, 2019 at 2:55 am #

REPLY

Hi, thanks for this helpful article. I'm a beginner and I need to do documents classification. In my model I use :

```
training_pipeline_data = [
('vectorizer', _create_vectorizer(lang)),
('densifier', _create_densifier()),
('scaler', _create_scaler()),
('classifier', _create_classifier())
]
training_pipeline = ibpip.Pipeline(training_pipeline_data)
                    training_pipeline.fit(features, labels)
```

training_pipeline.fit(features, labels)

with

```
def _create_vectorizer(language):
 stop_words = safe_get_stop_words(language) if language != 'en' else 'english'
 return TfidfVectorizer(sublinear_tf=True, min_df=7, norm='l2', ngram_range=(1, 2),
 encoding='latin-1', max_features=500, analyzer='word',
                        stop_words=stop_words)

def _create_densifier():
    return FunctionTransformer(lambda x: x.todense(), accept_sparse=True, validate=False)

def _create_scaler():
    return StandardScaler()

def _create_classifier():
    return GradientBoostingClassifier(n_estimators=160, max_depth=8, random_state=0)
```

Do I have to save in the pickel file the whole pipeline or just the classifier ?

When I save the whole pipeline, the size of the pickel file increases with the amount of training data, but I thouht it shouldn't impact the model size (only the parameters of the model should impact the size of this one)

REPLY

Jason Brownlee May 18, 2019 at 7:40 am #

Pickle all of it.

Yes, you are saving the mapping of words to numbers, it includes the whole known vocab required to encode new samples in the future.

REPLY

Krtin Ahuja June 17, 2019 at 9:54 pm #

Hi Jason,
How can I load the model to predict further?

REPLY

Jason Brownlee June 18, 2019 at 6:39 am #

I show how to load the model in the above tutorial.

What problem are you having exactly?

REPLY

Raphael June 21, 2019 at 8:00 pm #

Hi, big fan of your tutorials.
What are you thought about ONNX (https://onnx.ai/)
Do you think about making a tutorial to explain how it works and how to use it ?

Jason Brownlee June 22, 2019 at 6:38 am #

REPLY

What is ONNX? Perhaps you can summarize it for me?

Constantine June 24, 2019 at 1:43 am #

REPLY

Hi, thanks for the very useful post, as always!

I wanted to ask you, does this procedure work for saving Grid – Searched models as well? Because when I try to save the grid-search.best_estimator_ it does not give me the results I expect it to (ie the same score on the sample data I use) and the solutions I have found don't work either. Any tips on how to do that?

Many thanks!

Jason Brownlee June 24, 2019 at 6:35 am #

REPLY

Typically we discard grid search models as we are only interested the configuration so we can fit a new final model.

Constantine June 24, 2019 at 5:26 pm #

REPLY

Could you please point me to a source which shows how this is done in code? I 've tried (via my search) the following and it does not give me the expected results:

```
grid_elastic = GridSearchCV(elastic, param_grid_elastic,
 cv=tscv.split(X),scoring='neg_mean_absolute_error', verbose=1)
 grid_elastic.fit(X,y)
 print(grid_elastic.score(X,y))
 filename = 'finalized_model_grid.sav'
 joblib.dump(grid_elastic.best_params_, filename,compress=1)
 loaded_params_grid = joblib.load(filename)
 elastic = ElNet().set_params(**loaded_params_grid)
 elastic.fit(X,y)
 result = elastic.score(X, y)
 print(result)
```

I grid search an example model, fit it, calculate an example metric for comparisons, and then attempt to save the parameters and use them to instantiate the best estimator later, to avoid having to redo the exhaustive search. It does not give me the same score though. What is wrong? I 've tried just saving the best_estimator_ but it gives me the same wrong result.

Jason Brownlee June 25, 2019 at 6:14 am #

REPLY

Great question, and this is very common.

Machine learning algorithms are stochastic and we must average their performance over multiple runs.

You can learn more here:
https://machinelearningmastery.com/faq/single-faq/why-do-i-get-different-results-each-time-i-run-the-code

REPLY

ishrat July 3, 2019 at 9:59 pm #

firstly, thank you for sharing such amazing information always.
this is my code:

```
import time
 import numpy as np
 import pandas as pd
 from nltk import word_tokenize
 from nltk import pos_tag
 from nltk.corpus import stopwords
 from nltk.stem import WordNetLemmatizer
 from sklearn.preprocessing import LabelEncoder
 from collections import defaultdict
 from nltk.corpus import wordnet as wn
 from sklearn.feature_extraction.text import TfidfVectorizer
 from sklearn import model_selection, svm
 from sklearn.metrics import accuracy_score
 from sklearn.ensemble import RandomForestClassifier
 import pickle

start_time = time.time()
 np.random.seed(500)

#getting only the required columns and rows
 dataset = pd.read_csv("records.csv", sep="\t")
 dataset_new = dataset.iloc[:, [4, 5, 6, 8, 9]]

df = dataset_new.dropna(subset=['Debit'])
 df_required = df.iloc[:, [0, 2]]
 df_required = df_required[df_required['Description'] != 'OPENING BALANCE']
 df_less = df_required.iloc[:,:]
 df_less = df_less.reset_index(drop=True)

# dataset cleanup
 df_less = df_less.dropna(subset=['Description'])
 df_less = df_less.dropna(subset=['First Level Category'])
 df_less['description'] = " "

for index, row in df_less.iterrows():
 row['description'] = row['Description'].replace("-", " ")
 row['description'] = row['description'].replace("/", " ")
```

```
 row['description'] = row['description'].replace("_", " ")
 row['description'] = row['description'].replace(":", " ")
    row['description'] = row['description'].replace(".", " ")

dataset_time = time.time()
 print ("Time taken to create dataset : ", dataset_time – start_time)

df_less['description'] = [entry.lower() for entry in df_less['description']]
 df_less['description'] = [word_tokenize(entry) for entry in df_less['description']]
 df_less = df_less.reset_index(drop=True)

tokenize_time = time.time()
 print ("Time taken to tokenize dataset : ", tokenize_time – dataset_time)

for index, entry in enumerate(df_less['description']):
 Final_words = []
 for word in entry:
 if word.isalpha():
 Final_words.append(word)
    df_less.loc[index, 'desc_final'] = str(Final_words)

df_others = df_less[df_less['desc_final'] == '[]']
 df_less_final = pd.DataFrame()
 df_less_final = df_less[df_less['desc_final'] != '[]']
 df_less_final = df_less_final.reset_index(drop=True)
 data_cleanup_time = time.time()
 print ("Time taken for data cleanup", data_cleanup_time – tokenize_time)

Train_X, Test_X, Train_Y, Test_Y = model_selection.train_test_split(df_less_final['desc_final'],
 df_less_final['First Level Category'], test_size=0.33,
                                                    random_state=10,shuffle=True)

Tfidf_vect = TfidfVectorizer(max_features=106481)
 Tfidf_vect.fit(df_less['desc_final'])
 Train_X_Tfidf = Tfidf_vect.transform(Train_X)
 Test_X_Tfidf = Tfidf_vect.transform(Test_X)

clf = RandomForestClassifier(n_jobs=8, random_state=10)

clf.fit(Train_X_Tfidf, Train_Y)
 RandomForestClassifier(bootstrap=True, class_weight=None, criterion='gini',
 max_depth=None, max_features='auto', max_leaf_nodes=None,
 min_impurity_split=1e-07, min_samples_leaf=20,
 min_samples_split=2, min_weight_fraction_leaf=0.0,
 n_estimators=100, n_jobs=8, oob_score=False, random_state=10,
 verbose=0, warm_start=False)
 preds = clf.predict(Test_X_Tfidf)
 print("Random forest Accuracy Score -> ", accuracy_score(preds, Test_Y) * 100)
 preds.tofile("foo.csv", sep = '\n')
 dff = pd.DataFrame()
 col_name = ['category']
 dff = pd.read_csv("foo.csv", names = col_name, sep="\n")
```

sir this a model that i have prepared now i want to dump it using pickle but i am not able to understand how can i do this…since everytime i want to predict new records i want to preprocess my one of my rows as i am doing above, and also used vectorizer, and then predict the results, can

my one of my rows as I am doing above, and also used vectorizer..and then predict the results..can you please help me with the solution.
thank you

REPLY

Jason Brownlee July 4, 2019 at 7:47 am #

This is a common question that I answer here:
https://machinelearningmastery.com/faq/single-faq/can-you-read-review-or-debug-my-code

REPLY

teimoor July 17, 2019 at 4:14 am #

hi how can i learn python fast for the purpose of deep learning models like lstm ? can you notify me on gmail please

REPLY

Jason Brownlee July 17, 2019 at 8:30 am #

Right here:
http://machinelearningmastery.com/crash-course-python-machine-learning-developers/

REPLY

Raghad July 22, 2019 at 5:59 pm #

Thank you so much for all your effort, but I am a beginner in ML and Python and I have a basic conceptual question:
I used a CSV file to train, test and fit my random forest model then I saved the model in a pickle file. Now my partner wants to use the model for prediction on new unseen data(entered by user) so my question is should I send her only the model I saved in a pickle file or also the data I used to train and fit the model? I mean would the pkl model work even if the CSV file containing the data used to fit the model is not in the same folder or host? I hope my question is clear. Thank you again very much!!

REPLY

Jason Brownlee July 23, 2019 at 7:57 am #

Just the model is required.

REPLY

Chandan Kumar Jha August 20, 2019 at 3:05 am #

Sir, model saving and re-using is okay but what about the pre-processing steps that someone would have used like LabelEncoder or StandardScalar function to transform the features.

We would need to apply the same transformation on the unseen dataset so how do we

proceed there? How can we save these pre-processing steps.

REPLY

Jason Brownlee August 20, 2019 at 6:28 am #

Yes, pre-processing must be identical. You might want to save the objects involved.

REPLY

Fathima August 31, 2019 at 2:35 am #

I have trained the model using python 3.7, will i be able to test it using python 3.5?

REPLY

Jason Brownlee August 31, 2019 at 6:11 am #

Not sure, perhaps test it and see?

REPLY

Ned H September 8, 2019 at 11:43 pm #

Hi Jason,

I'm a big fan of your blog. The thing is, while it is useful to save a model, often the model is already part of a pipeline. I've had success using the joblib method to store a pre-trained pipeline and then load it into the same environment that I've built it in and get predictions. However, when I say, save a pipeline in AWS and then load it locally, I get errors.

Is there a best practice when it comes to saving pipelines vs naked models?

Perhaps a tutorial where you train a pipeline using RandomizedSearchCV and then save it would be useful?

Thanks for all your great tutorials!

REPLY

Jason Brownlee September 9, 2019 at 5:16 am #

What errors do you get?

Saving/loading a pipeline is the same as saving/loading a single model as far as I understand.

REPLY

Alban October 2, 2019 at 1:48 am #

Hi Jason, thanks for your time, and really interesting tutorials !
I trained and saved a random forest model, and i analysed the classification performance with different thresholds.

Now i want to apply this saved random forest with a new data set, to get predictions, but using a different threshold than 50%. But it seems i can't get the outcome of rf.predict_proba(x) function, i get a "NotFittedError"… it says that my rf model is not fitted yet… i am lost now… Is there sthg wrong in my reasoning ? Is there an other way to get the classification probabilities ?Thank you.

REPLY

Jason Brownlee October 2, 2019 at 8:02 am #

Perhaps confirm the model was fit, and the fit model was saved?

REPLY

Rahel October 16, 2019 at 6:00 pm #

Hi Jason..there is a error in line number 13 of the code…instead of "# Fit the model on 33%" it should be "# Fit the model on 67%" as we are fitting the model to the training set which is 67%…

REPLY

Jason Brownlee October 17, 2019 at 6:25 am #

Thanks, fixed.

REPLY

Shiva Vutukuri November 6, 2019 at 3:55 am #

Hi Jason,
I am working on APS failure scania trucks project…

After using joblib library for saving & loading the file, i got the following:

```
# Save model for later use
 modelName = 'finalModel_BinaryClass.sav'
 joblib.dump(finalModel, modelName)
 Output: ['finalModel_BinaryClass.sav']

# load the model from disk
 loaded_model = joblib.load(modelName)
 result = loaded_model.score(X_validation, Y_validation)
 print(result)
 Output: 0.9894375
```

My query is i am unable to find where the final model is saved… Could you please help me?

REPLY

Jason Brownlee November 6, 2019 at 6:45 am #

It will be in your current working directly, e.g. where you are running the code.

Shiva Vutukuri November 6, 2019 at 8:18 pm #

REPLY

Thanks a lot Mr. Jason

Jason Brownlee November 7, 2019 at 6:40 am #

REPLY

You're welcome.

KK December 9, 2019 at 3:44 pm #

REPLY

Hi Sir,

Excellent tutorial.

Could you please tell me, why you used .sav format to save the model?

Can we use .pkl format instead. What is the advantage of .sav over .pkl or any other format.

Thank you,
KK

Jason Brownlee December 10, 2019 at 7:25 am #

REPLY

We use the pickle format in this tutorial. You can use any file extension you wish.

pat c December 13, 2019 at 9:47 am #

REPLY

if you build a model using class weights, do you need to account for that in any way when scoring a new dataset?

Jason Brownlee December 13, 2019 at 1:41 pm #

REPLY

No, but you should select a metric that best captures what is important about the predictions.

E.g. use F-measure or G-mean, or precision, ROC AUC, etc.

vikash December 30, 2019 at 12:40 am #

REPLY

Hi,
I have also used Standardization on the training and testing dataset. I have deployed the model file. How can I do standardization when calling the model through API?

Jason Brownlee December 30, 2019 at 6:00 am #

REPLY

This will help:
https://machinelearningmastery.com/how-to-save-and-load-models-and-data-preparation-in-scikit-learn-for-later-use/

SHubham January 9, 2020 at 4:30 pm #

REPLY

Hi Jason,

Can you please tell me how to convert a .pkl file to .pb file or .tflite file?

Jason Brownlee January 10, 2020 at 7:23 am #

REPLY

I don't know off hand, perhaps try posting to stackoverflow?

Murat Kulustepe January 13, 2020 at 1:55 am #

REPLY

Hi Jason !

My model was created with exel file..Anyway model is ready. Now I would like to use model online. I mean ınputs are will come from sql database and same time I would like to see result from model. (Predıct value)

Can you explain me please how to do ?

Jason Brownlee January 13, 2020 at 8:28 am #

REPLY

Yes, implement your model in code, then pass one row at a time to your model to make predictions.

Murat Kulustepe January 13, 2020 at 9:16 am #

REPLY

Ok but how ? I am asking because I am not sure how to do?
If you can explain me a little bit, or you can Show me example (example link) I would be happy.!

Thanks so much.

Jason Brownlee January 13, 2020 at 1:42 pm #

REPLY

I don't know how to answer you – the question is two broad. You have complete freedom over how you code your own algorithm and save it.

Perhaps you can narrow down your question. Also, I don't have the capacity to implement your algorithm for you.

If you prefer to use a library, the example in the above tutorial would be a good start.

josheeg January 20, 2020 at 1:24 am #

REPLY

Can these model files pickles be opened or created in anything else?
I know orange uses some of scikit learn and ran both.
But I never made a scikit learn pickle and opened it in orange or created a orange save model wiget file is a pickle file.
I would like to use this to open a model in keras tensorflow and be able to run the apps to make it tensorflow light compatible and then run it in a microcontroller. Like rasbery pi 4 or maby the requirements is it has to run python 3 there are some arm processors that do that.

Jason Brownlee January 20, 2020 at 8:41 am #

REPLY

Perhaps.

You cannot pickle a keras model, you must use the h5 format:
https://machinelearningmastery.com/save-load-keras-deep-learning-models/

Pierre January 20, 2020 at 6:46 am #

REPLY

Hi Jason, I was working through this from your ML Master w/ Python Book and ran into this error:

Traceback (most recent call last):
File "/Users/pierrenoujeim/Desktop/MLDS/Python/MasterML/ml w: python/code/17. Save and Load Models/pickletest.py", line 2, in
import pandas
File "/anaconda3/lib/python3.6/site-packages/pandas/__init__.py", line 19, in
"Missing required dependencies {0}".format(missing_dependencies))
ImportError: Missing required dependencies ['numpy']

I copied the code as is from this page and got the same error. What should I do?

Jason Brownlee January 20, 2020 at 8:46 am #

REPLY

Ouch. I have not seen that before.

Perhaps confirm that Python and scipy are installed correctly:

https://machinelearningmastery.com/setup-python-environment-machine-learning-deep-learning-anaconda/

And if so, perhaps search or post the error to stackoverflow.

Manuel January 20, 2020 at 12:52 pm #

REPLY

Jason one question… how we can export the results of our model to Excel or ASCII file?

thanks

Jason Brownlee January 20, 2020 at 2:07 pm #

REPLY

See this tutorial:
https://machinelearningmastery.com/how-to-save-a-numpy-array-to-file-for-machine-learning/

Madan Kumar Y January 28, 2020 at 5:35 pm #

REPLY

Hi Jason,
How can i unpickle the learnable parameters(weights and biases) after Fitting the model. I am finding hard to get the learnable parameters from the pickle file. Is there any process??
Thanks

Jason Brownlee January 29, 2020 at 6:30 am #

REPLY

Each model stores its internal parameters differently.

The sckit-learn API explains how to access the parameters of each model, once loaded.

hayder February 12, 2020 at 9:15 pm #

REPLY

how can I store the output of one class svm to buffer in python?

Jason Brownlee February 13, 2020 at 5:38 am #

REPLY

See this:
https://machinelearningmastery.com/how-to-save-a-numpy-array-to-file-for-machine-learning/

Jose Q March 22, 2020 at 1:59 pm #

REPLY

Hi Jason!

Thank you for this tutorial! You always explain concepts so easy!

I have a question.

If I had to use a scaler during training like

…

X_scaled = scaler.fit_transform(X)

…

Later I fitted the model and saved it with pickle.dump(clf, open(filename, ‘wb’))

Then I supone that I have to scale test data for prediction using the same scaler fitted during training like scaler.transform(x_test)

My question is: besides saving the model, do we have to save objects like the scaler in this example to provide consistency? Can we use pickle or joblib ?

Thank you

REPLY

Jason Brownlee March 23, 2020 at 6:11 am #

Thanks.

Yes, save the model and any data prep objects, here is an example:
https://machinelearningmastery.com/how-to-save-and-load-models-and-data-preparation-in-scikit-learn-for-later-use/

REPLY

Jose Q March 23, 2020 at 7:17 am #

Thank you! You rock !

REPLY

Jason Brownlee March 23, 2020 at 7:47 am #

Thanks!

REPLY

Akilu Rilwan March 24, 2020 at 11:21 pm #

Hi Jason,

Thank you for all the pieces you put here. I always find your resources very useful.

Many thanks.

REPLY

Jason Brownlee March 25, 2020 at 6:33 am #

Thanks!

Thanks!

Adarsh April 11, 2020 at 12:56 am #

REPLY

Hi Jason,
I am using Django to deploy my model to the web….. I am getting a strange behaviour with the labelencoder of my model…..Can you please help me to go through this by posting an article:
How to deploy a sklearn model in django..
please
Best Regards,
Adarsh

Jason Brownlee April 11, 2020 at 6:23 am #

REPLY

No idea, sorry. Perhaps try posting on stackoverflow.

AnanShekher Srivastava April 17, 2020 at 11:05 pm #

REPLY

Thank you. This post shows how to save a model once after being trained on the entire dataset in one go. However, most of the real-world data sets are huge and can't be trained in one go. How can I save a model after training it on each chunk of data?

```
df = pd.read_csv("an.csv", chunksize=6953)
 for chunk in df:
 text = chunk['body']
    label = chunk['user_id']

   X_train, X_test, y_train, y_test = train_test_split(text, label, test_size=0.3 )

   text_clf = Pipeline([('vect', TfidfVectorizer()),
('tfidf', TfidfTransformer()),
('clf', LinearSVC()),
                        ])

   text_clf.fit(X_train, y_train)

   # save the model to disk
 filename = 'finalized_model.sav'
    joblib.dump(model, filename)
```

Saving it this way will give me the model trained on the last chunk. I want the model trained on every chunk. Any help?

Jason Brownlee April 18, 2020 at 5:58 am #

REPLY

Really, I don't agree.

If you can't fit your data in memory, perhaps look into using a framework like hadoop with mahout to fit models using online methods?

REPLY

AnanShekher Srivastava April 19, 2020 at 6:39 pm #

So are you saying saving that way will give me a model based on every chunk?

REPLY

Jason Brownlee April 20, 2020 at 5:26 am #

No, there are algorithms and versions of algorithms that support iterative learning algorithms – called online learning.

REPLY

Michael Nguyen April 18, 2020 at 2:16 am #

Hi Jason,

i forgot parameter of saved model. How can i get parameter of model from .sav file?

REPLY

Jason Brownlee April 18, 2020 at 6:06 am #

Most sklearn models store the parameters used to configure their instance as a property (I think…). Check the contents of the loaded object, or check the sklearn api.

REPLY

Martino Innocenti May 14, 2020 at 7:15 pm #

Hi,
is possible to convert a sklearn model in tensorflowlite model (.tflite)?
i need to run an SVM model in android and this seems to me the best solution (if it is possible)
are there other solution?

REPLY

Jason Brownlee May 15, 2020 at 5:57 am #

I don't know, sorry.

REPLY

Martin M. May 19, 2020 at 6:41 pm #

Great post, thank you!

I tried this out but noticed that every time I serialize a model, the resulting file has a different checksum. Any ideas how to preserve the integrity of the model file if data and algorithms do not

checksum. Any ideas how to preserve the integrity of the model file if data and algorithms do not change?

Here is my test code:

```
import numpy as np
from sklearn import linear_model
import joblib
import hashlib

# set a fixed seed …
np.random.seed(1979)

# internal md5sum function
def md5(fname):
hash_md5 = hashlib.md5()
with open(fname, “rb”) as f:
for chunk in iter(lambda: f.read(4096), b””):
hash_md5.update(chunk)
    return hash_md5.hexdigest()

# dummy regression data
X = [[0., 0., 0.,1.], [1.,0.,0.,0.], [2.,2.,0.,1.], [2.,5.,1.,0.]]
Y = [[0.1, -0.2], [0.9, 1.1], [6.2, 5.9], [11.9, 12.3]]

# create model
reg = linear_model.LinearRegression()

# save model to disk to make it persistent
with open(“reg.joblib”, “w”):
    joblib.dump(reg, “reg.joblib”)

# load persistant model from disk
with open(“reg.joblib”, “r”):
    model = joblib.load(“reg.joblib”)

# fit & predict
reg.fit(X,Y)
model.fit(X,Y)
myprediction1 = reg.predict([[2., 2., 0.1, 1.1]])
myprediction2 = model.predict([[2., 2., 0.1, 1.1]])

# run several times … why does md5sum change everytime?
print(md5(“reg.joblib”))
print(myprediction1, myprediction2)
```

Jason Brownlee May 20, 2020 at 6:23 am #

REPLY

Thanks!

The model will be different each time you train it, in turn different weights are saved to file.

Martin M. June 5, 2020 at 4:45 pm #

REPLY

Just to properly close this example and after some more investigation I can say the problem in my example stems from the joblib.dump serialization. The simple linear regression model with its weights is reproducible. For anybody interested, I tried to answer it here giving more context: https://stackoverflow.com/questions/61877496/how-to-ensure-persistent-sklearn-models-on-bit-level

Jason Brownlee June 6, 2020 at 7:45 am #

REPLY

Thanks for sharing!

Saket Nandan May 30, 2020 at 3:21 am #

REPLY

```
xgb_clf =xgb.XGBClassifier(base_score=0.5, booster='gbtree', colsample_bylevel=1,
colsample_bynode=1, colsample_bytree=0.7, gamma=0.0, gpu_id=-1,
importance_type='gain', interaction_constraints='',
learning_rate=0.1, max_delta_step=0, max_depth=10,
min_child_weight=3, monotone_constraints='()',
n_estimators=100, n_jobs=0, num_parallel_tree=1,
objective='binary:logistic', random_state=50, reg_alpha=1.2,
reg_lambda=1.6, scale_pos_weight=1.0, subsample=0.9,
        tree_method='exact', validate_parameters=1, verbosity=None)

xgb_clf.fit(X1, y1)
```