Saving the Trained Model

Mike West

MACHINE LEARNING ENGINEER

Module Overview

Serialization

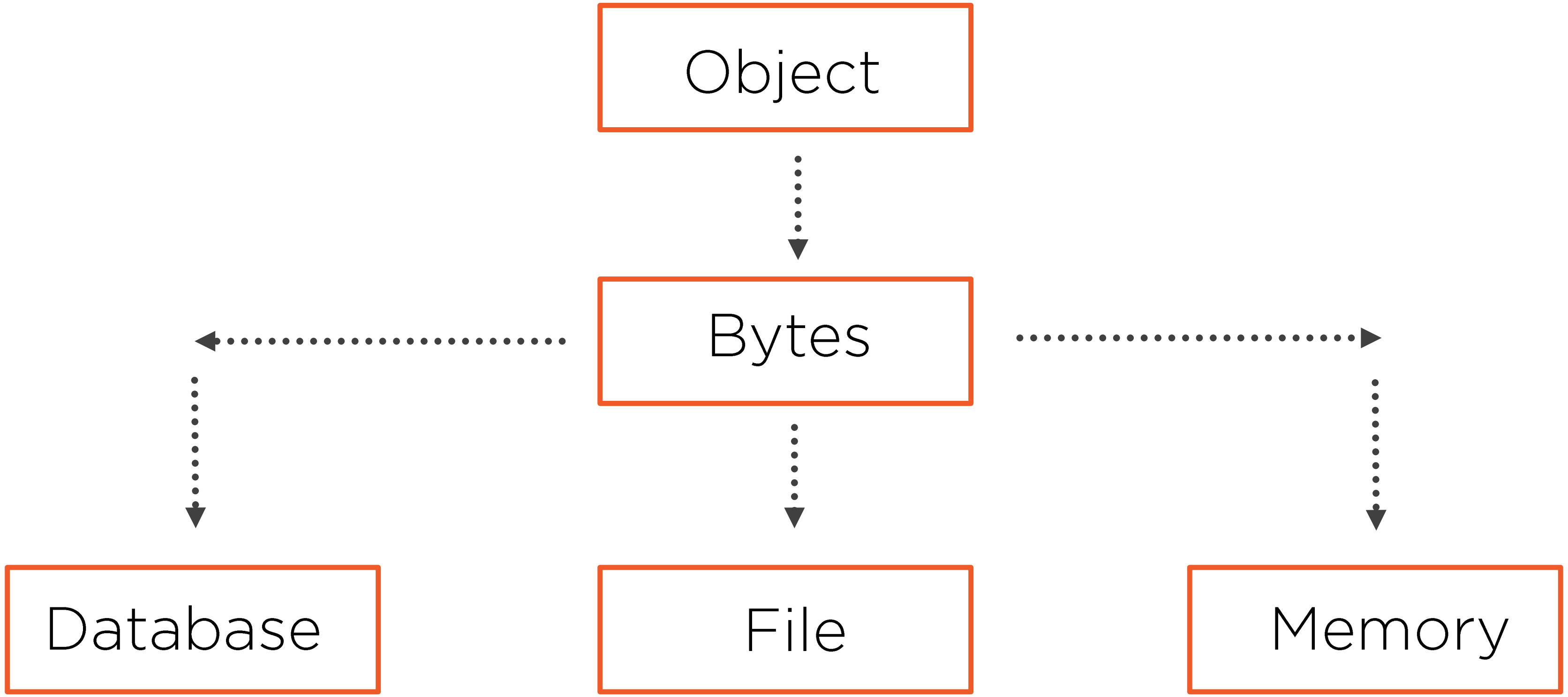
Pickle module

JSON

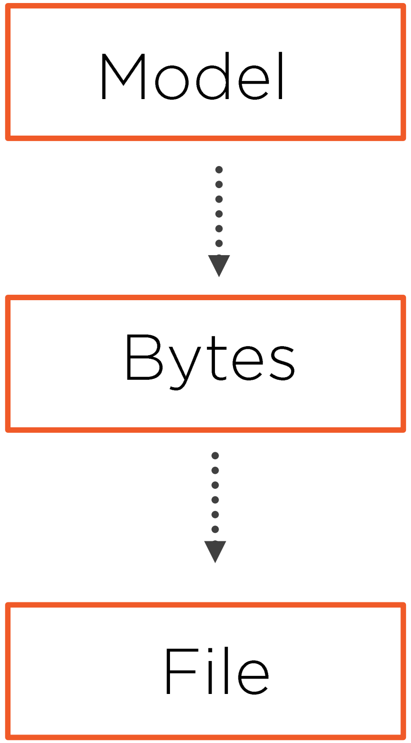
Pickle problems

Pickle and JSON demonstration

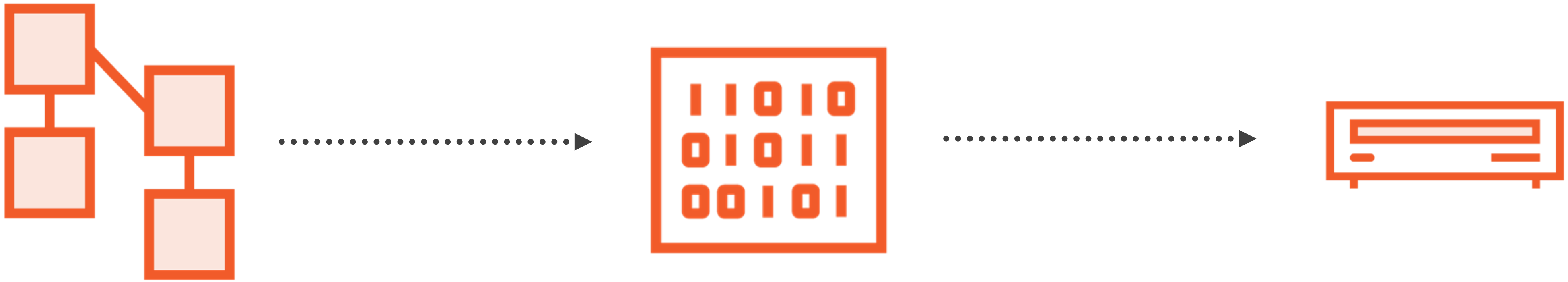
Serialization



Model Serialization



Pickle



Model

Byte Steam

Disk

Pickle Python Objects

Python Only

Pickle is protocol-specific to Python

Version Dependent

No guarantee of compatibility

Not Secure

Only unpickle data from a trusted source

```
{
  "species": "Dog",
  "breed": "Lab",
  "color": "Yellow",
  "age": 6
}
```

Key Value Pair in JSON

The two primary parts that make up JSON are keys and values

JSON to Python

```
import json

x = {
  "name": "John",
  "age": 30,
  "city": "New York"
     }

y = json.dumps(x)

print(y)
```

```
{"name": "John", "age": 30, "city": "New York"}
```

Pickle Problems

Pickling an object is the slowest approach to persisting your objects

No guarantee of compatibility between Python versions

Pickle is not secure against maliciously constructed data

Pickle is only for Python. All other languages need not apply

Demo

Import your libraries

Basic Pickle

Pickle models

JSON persistence

Saving XGBoost models

Summary

Defined serialization

Pickle

Persistence with JSON

Pickle problems

Saving XGBoost models