

Theoretical and practical metagenomic approaches to viral discovery

Practical Session: Practical Introduction to Machine Learning and SVMs

Kevin Lamkiewicz, Manja Marz

24.10.2019

European Virus Bioinformatics Center

WHOAMI

WHOAMI

- ▶ PhD student at Manja's Lab since 2017
- ▶ Interested in (RNA) viruses
- ▶ Focus on RNA secondary structures and ncRNAs
- ▶ Experience with Machine Learning, RNA structures, assemblies, alignments, SNP calling, ...



Introducing scikit-learn

BUILDING A SUPPORT VECTOR MACHINE WITH PYTHON3

```
1 #!/usr/bin/env python3
2
3 # to import scikit-learn we write the following
4 import sklearn
5
```

BUILDING A SUPPORT VECTOR MACHINE WITH PYTHON3

```
1 #!/usr/bin/env python3
2
3 # to import scikit-learn we write the following
4 import sklearn
5
6 # but actually we need this:
7 from sklearn import svm
8
```

BUILDING A SUPPORT VECTOR MACHINE WITH PYTHON3

```
1 #!/usr/bin/env python3
2
3 # to import scikit-learn we write the following
4 import sklearn
5
6 # but actually we need this:
7 from sklearn import svm
8
9 # now we create our SVM
10 classifier = svm.SVC(kernel='linear')
11
```

BUILDING A SUPPORT VECTOR MACHINE WITH PYTHON3

```
1 #!/usr/bin/env python3
2
3 # to import scikit-learn we write the following
4 import sklearn
5
6 # but actually we need this:
7 from sklearn import svm
8
9 # now we create our SVM
10 classifier = svm.SVC(kernel='linear')
11
12 # Done.
```

UNDERSTANDING THE TWO LINES OF CODE

```
1 from sklearn import svm
```

- ▶ make all functions of the `svm` class available

UNDERSTANDING THE TWO LINES OF CODE

```
1 from sklearn import svm
```

- ▶ make all functions of the `svm` class available

```
1 classifier = svm.SVC(kernel='linear')
```

- ▶ create a classifier object with a linear kernel

Applying machine learning to real biological data

THE IRIS TEST DATA SET

```
1 from sklearn import svm, datasets  
2 # load the iris dataset  
3 iris = datasets.load_iris()  
4 print(iris)
```

COFFEE BREAK

