MANCHESTER 1824
The University of Manchester

Part II

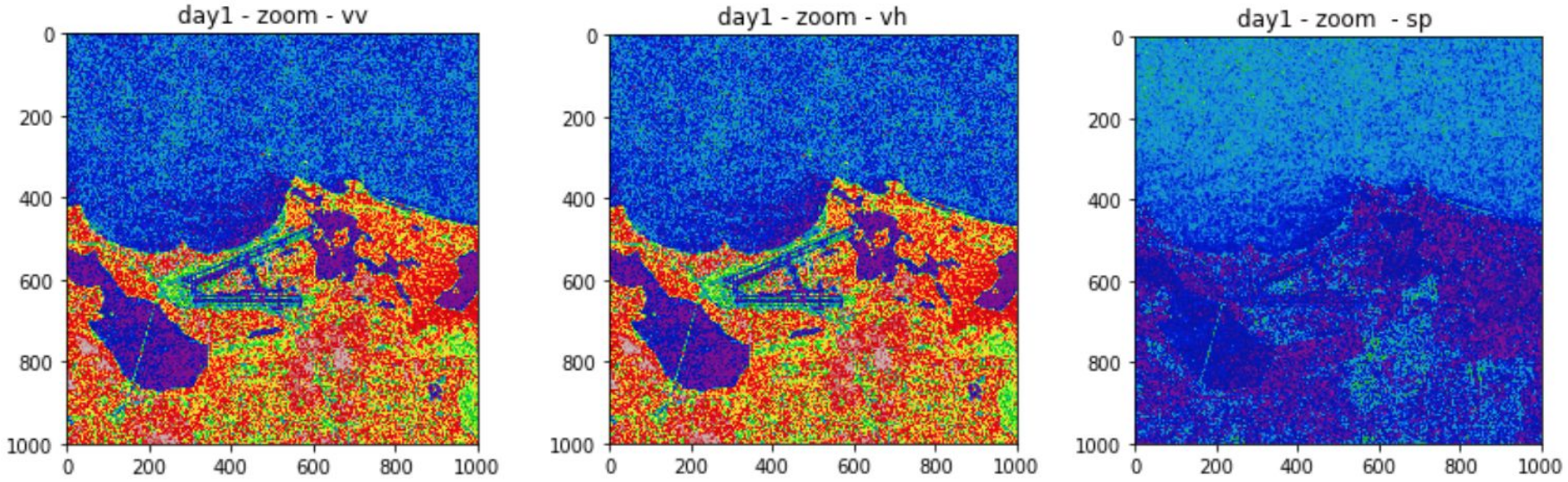
Dr Joe Fennell | Jodrell Bank Centre for Astrophysics, University of Manchester, UK | joseph.fennell@manchester.ac.uk

MANCHESTER 1824
The University of Manchester

Clustering: Code walkthrough

In [1]:
```
# Import needed modules

import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import glob
from skimage.segmentation import slic
from skimage.segmentation import mark_boundaries
```



Dr Joe Fennell | Jodrell Bank Centre for Astrophysics, University of Manchester, UK | joseph.fennell@manchester.ac.uk

MANCHESTER 1824
The University of Manchester

Clustering: Code walkthrough

In [6]:
```
number_segments=50000
segments = slic(day1_training, n_segments=number_segments, compactness=50, sigma = 0,convert2lab=F
alse)
plt.figure(figsize=(20,20))
plt.imshow(mark_boundaries(day1_training[:,:,0]/255, segments))
plt.show()
```

In [10]:
```
from sklearn import cluster,preprocessing

kmeans_3clusters=cluster.KMeans(n_clusters=4)
kmeans_3clusters.fit(seg_means)
```

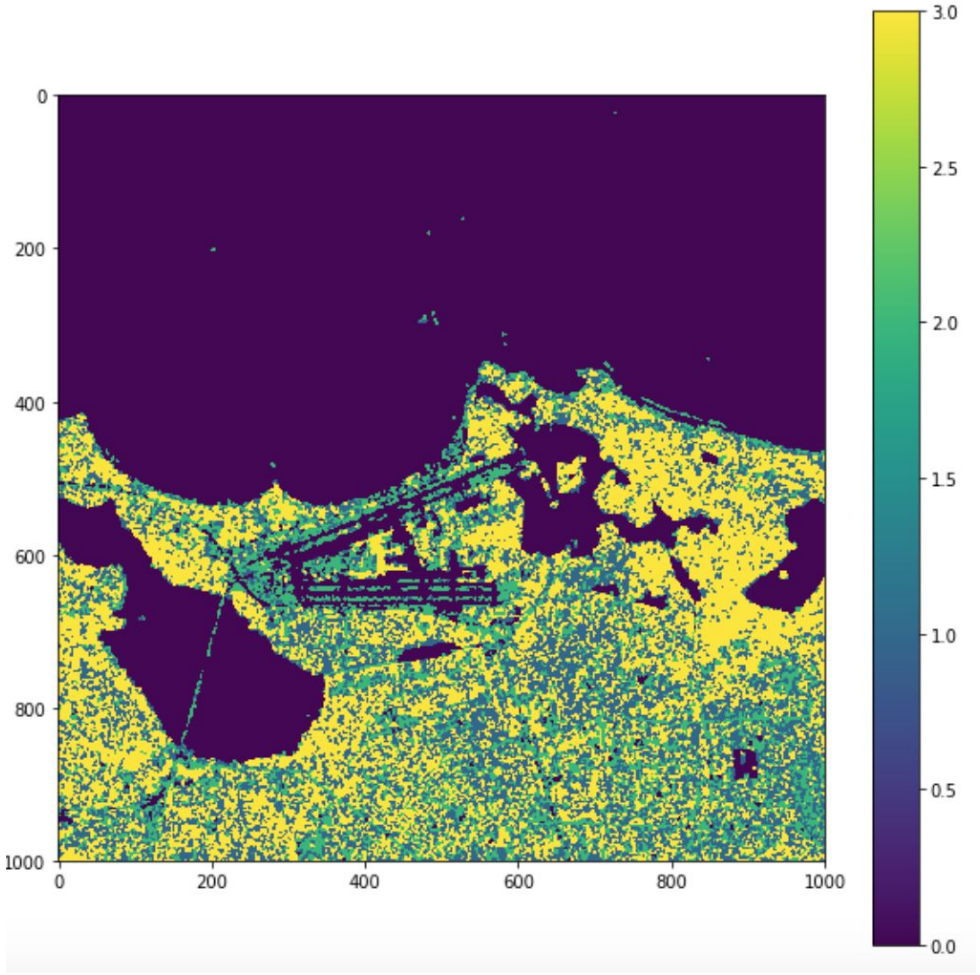
Out[10]:
```
KMeans(algorithm='auto', copy_x=True, init='k-means++', max_iter=300,
    n_clusters=4, n_init=10, n_jobs=1, precompute_distances='auto',
    random_state=None, tol=0.0001, verbose=0)
```

In [11]:
```
plt.figure(figsize=(10,10))
plt.imshow(kmeans_3clusters.labels_[segments])
plt.colorbar()
```

Out[11]:
```
<matplotlib.colorbar.Colorbar at 0x7fb5e80997f0>
```


MANCHESTER 1824
The University of Manchester

Clustering: Code walkthrough



Dr Joe Fennell | Jodrell Bank Centre for Astrophysics, University of Manchester, UK | joseph.fennell@manchester.ac.uk

MANCHESTER 1824
The University of Manchester

Choosing a kernel

The *kernel function* can be any of the following:

- linear: $\langle x, x' \rangle$.
- polynomial: $(\gamma \langle x, x' \rangle + r)^d$. d is specified by keyword `degree`, r by `coef0`.
- rbf: $\exp(-\gamma \|x - x'\|^2)$. γ is specified by keyword `gamma`, must be greater than 0.
- sigmoid ($\tanh(\gamma \langle x, x' \rangle + r)$), where r is specified by `coef0`.

MANCHESTER 1824
The University of Manchester

$$K(x,z) = \phi(x)^T \phi(z).$$

The Kernel Trick

- We can map the parameters to a higher dimensional feature space
- Polynomial kernel for non-linearly separable data https://www.youtube.com/watch?v=3liCbRZPrZA
- However we don't have to calculate all of the new features explicitly….
- This is important for dealing with mappings that yield infinite-dimensional feature space
- e.g.

$$K(x,z) = \exp\left(-\frac{||x-z||^2}{2\sigma^2}\right).$$

MANCHESTER 1824
The University of Manchester

Remember the Cost term?

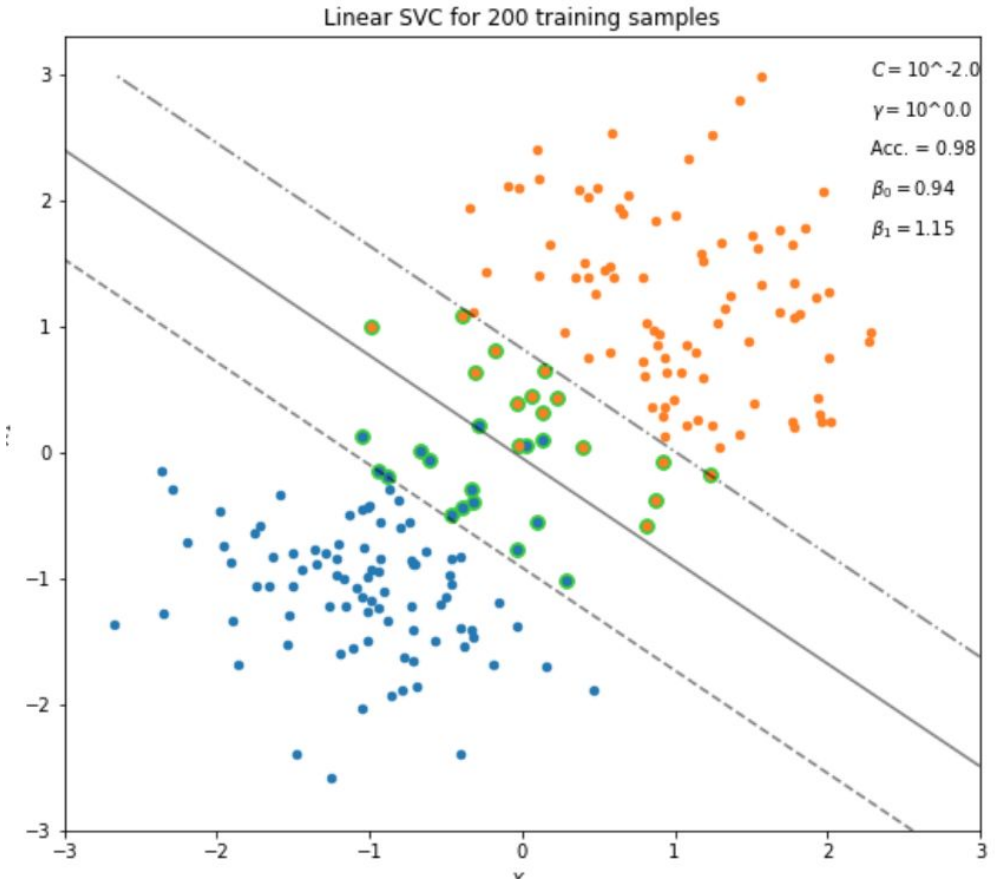
$$\min_{\gamma,w,b} \quad \frac{1}{2}||w||^2 + C\sum_{i=1}^{m}\xi_i$$

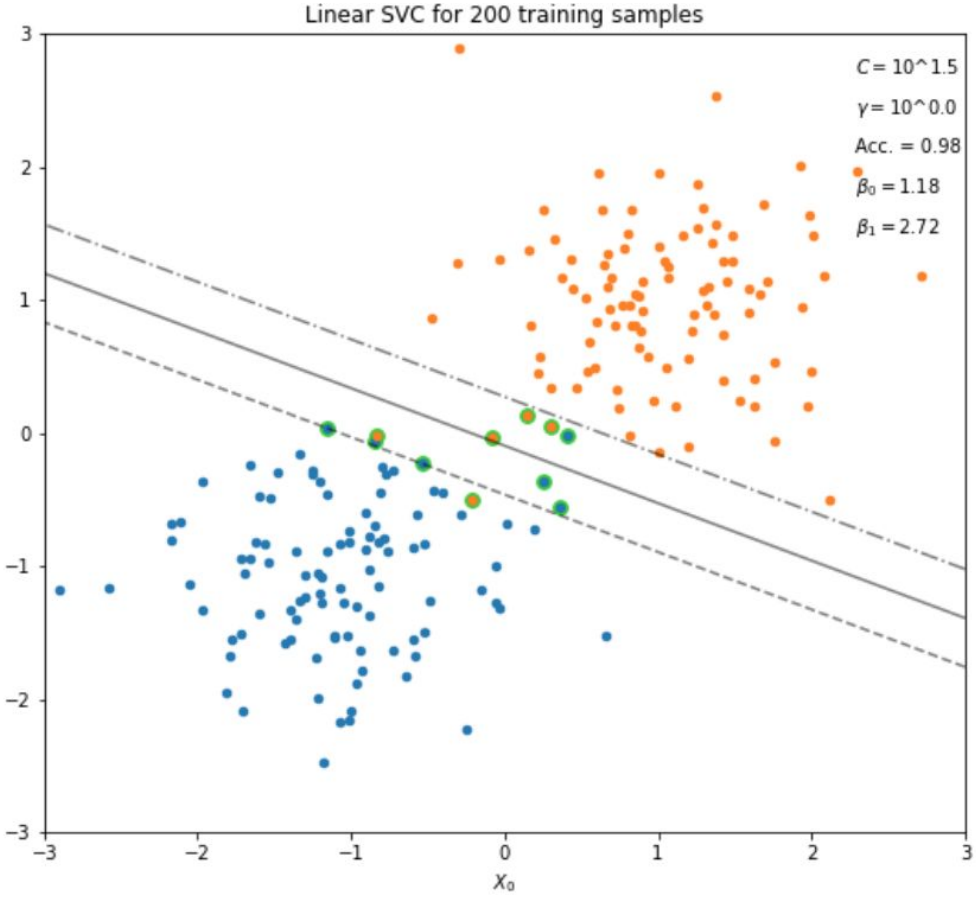
The cost term

$$\text{s.t.} \quad y^{(i)}(w^T x^{(i)} + b) \geq 1 - \xi_i, \quad i = 1, \ldots, m$$
$$\xi_i \geq 0, \quad i = 1, \ldots, m.$$

MANCHESTER 1824
The University of Manchester

Adjusting the C (penalty) term

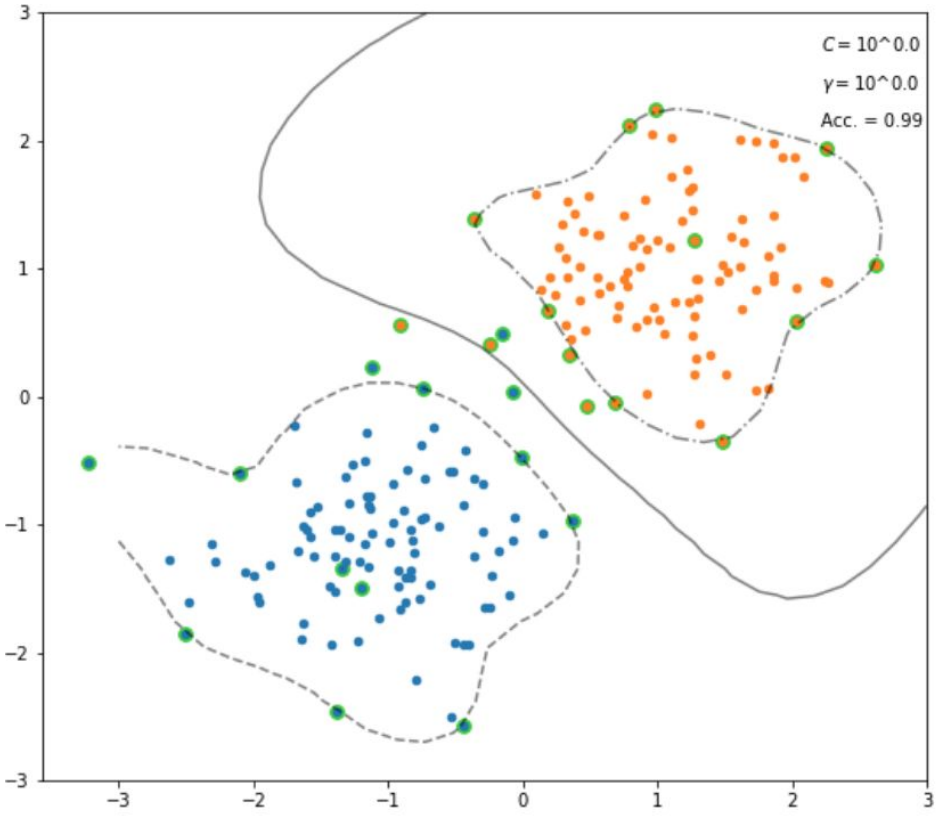


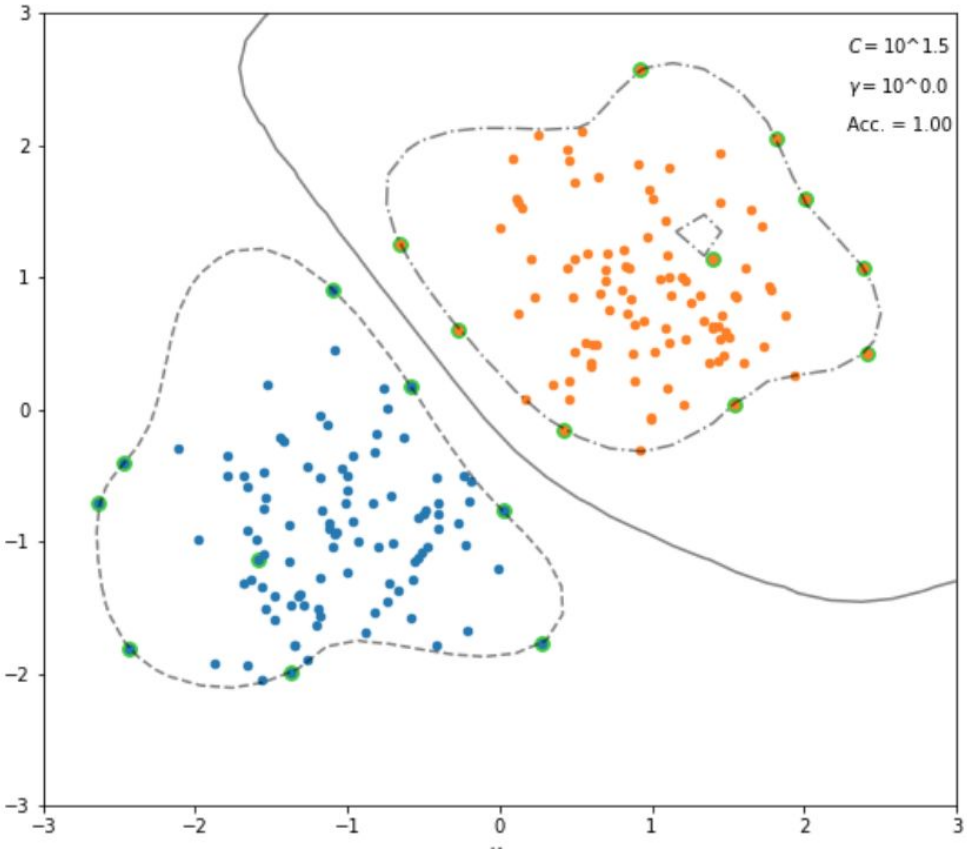


Dr Joe Fennell | Jodrell Bank Centre for Astrophysics, University of Manchester, UK | joseph.fennell@manchester.ac.uk

MANCHESTER 1824
The University of Manchester

Adjusting the Cost term (RBF)

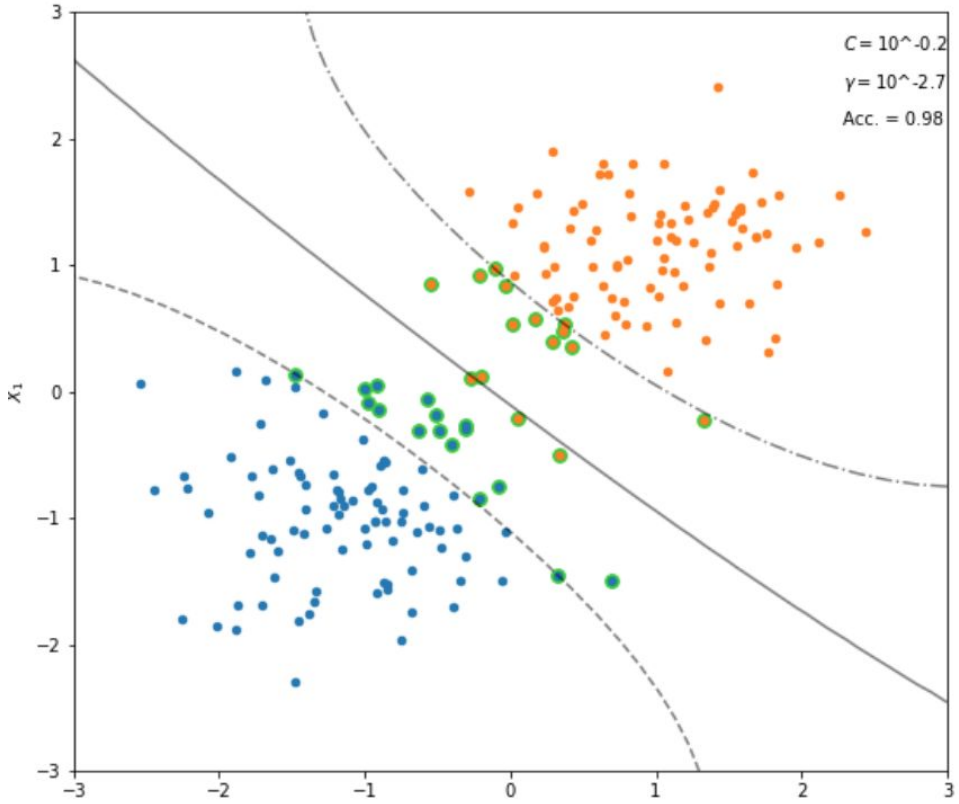


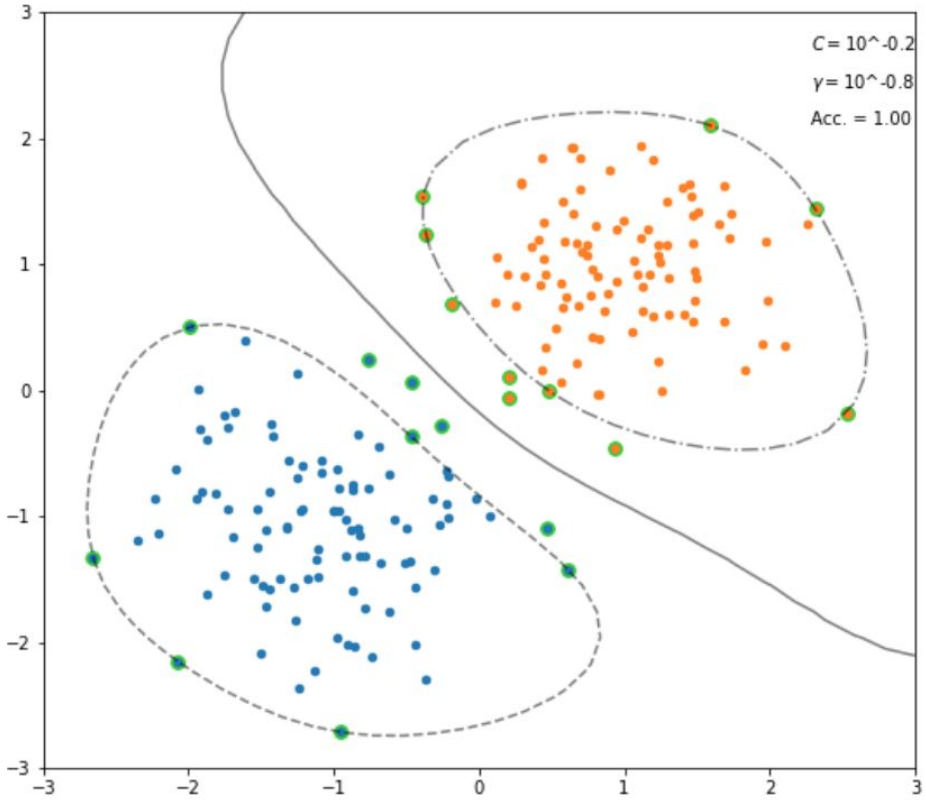


Dr Joe Fennell | Jodrell Bank Centre for Astrophysics, University of Manchester, UK | joseph.fennell@manchester.ac.uk

MANCHESTER 1824
The University of Manchester

Adjusting the gamma term



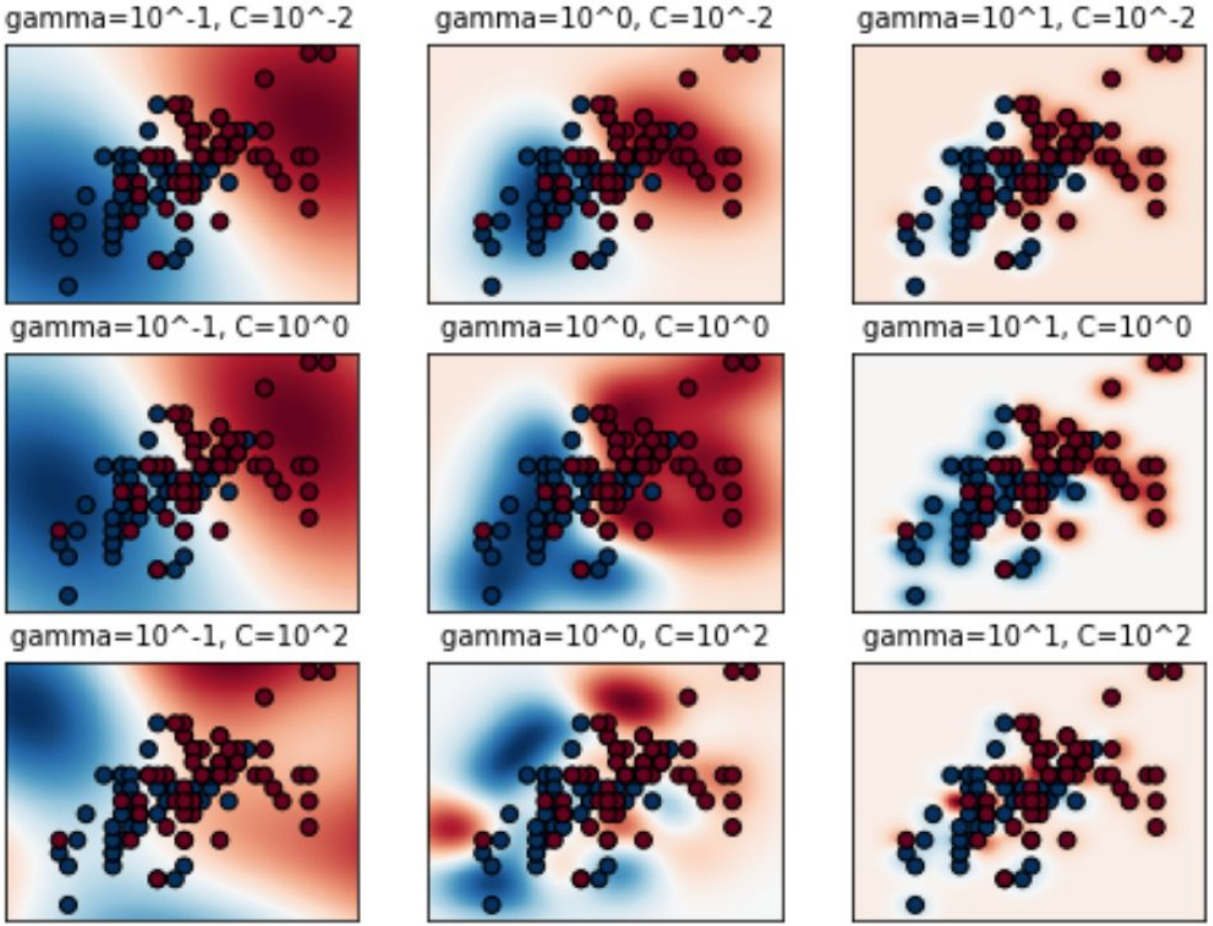


Dr Joe Fennell | Jodrell Bank Centre for Astrophysics, University of Manchester, UK | joseph.fennell@manchester.ac.uk

MANCHESTER 1824

The University of Manchester

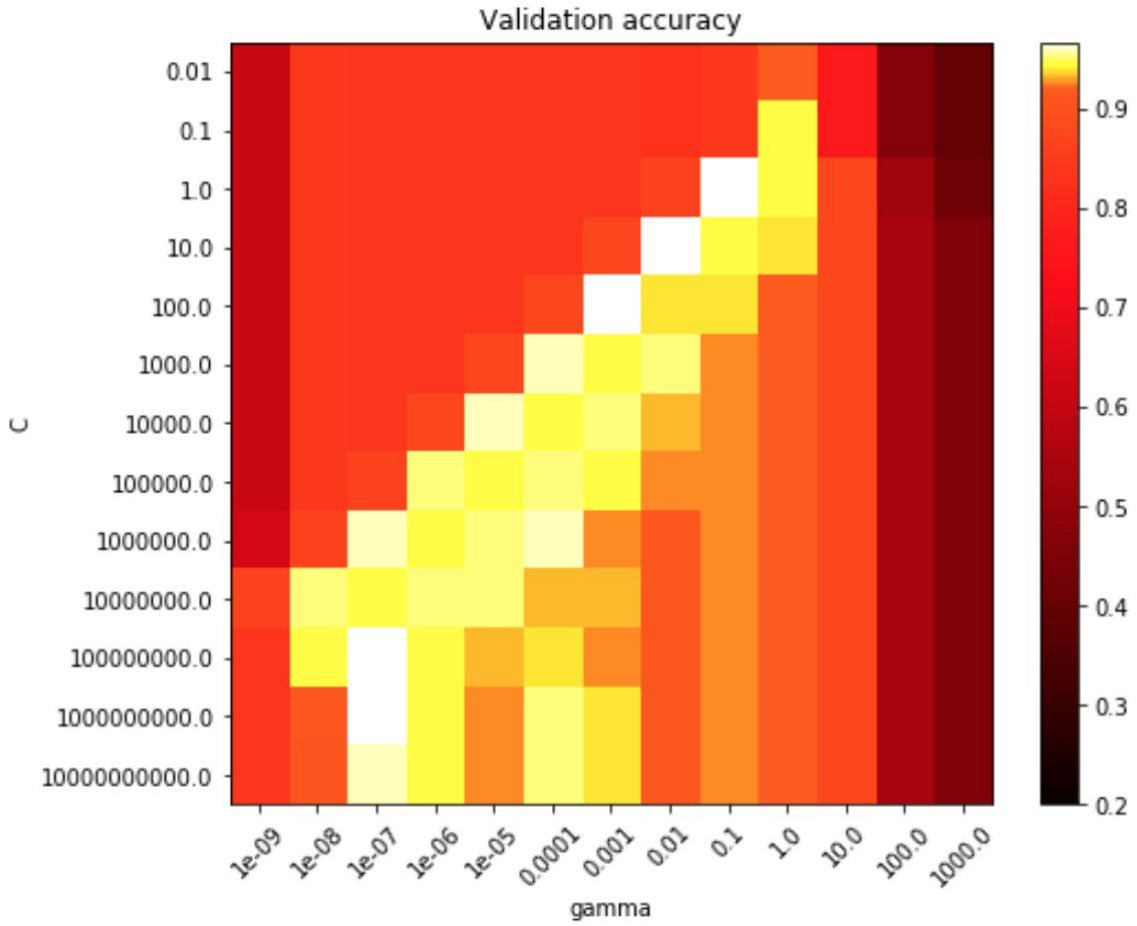
SVM hyperparameter tuning



Dr Joe Fennell | Jodrell Bank Centre for Astrophysics, University of Manchester, UK | joseph.fennell@manchester.ac.uk

MANCHESTER 1824
The University of Manchester

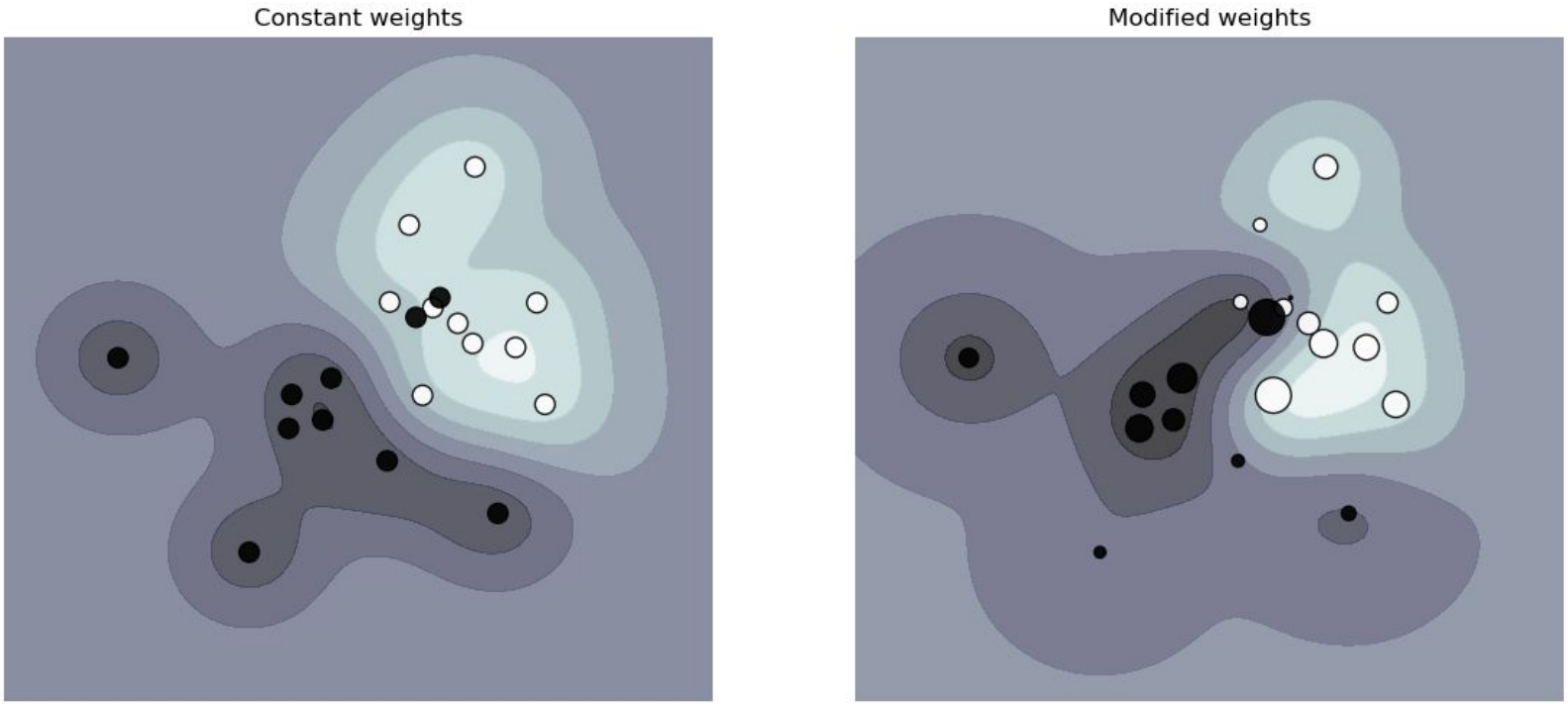
SVM hyperparameter tuning



Dr Joe Fennell | Jodrell Bank Centre for Astrophysics, University of Manchester, UK | joseph.fennell@manchester.ac.uk

MANCHESTER 1824
The University of Manchester

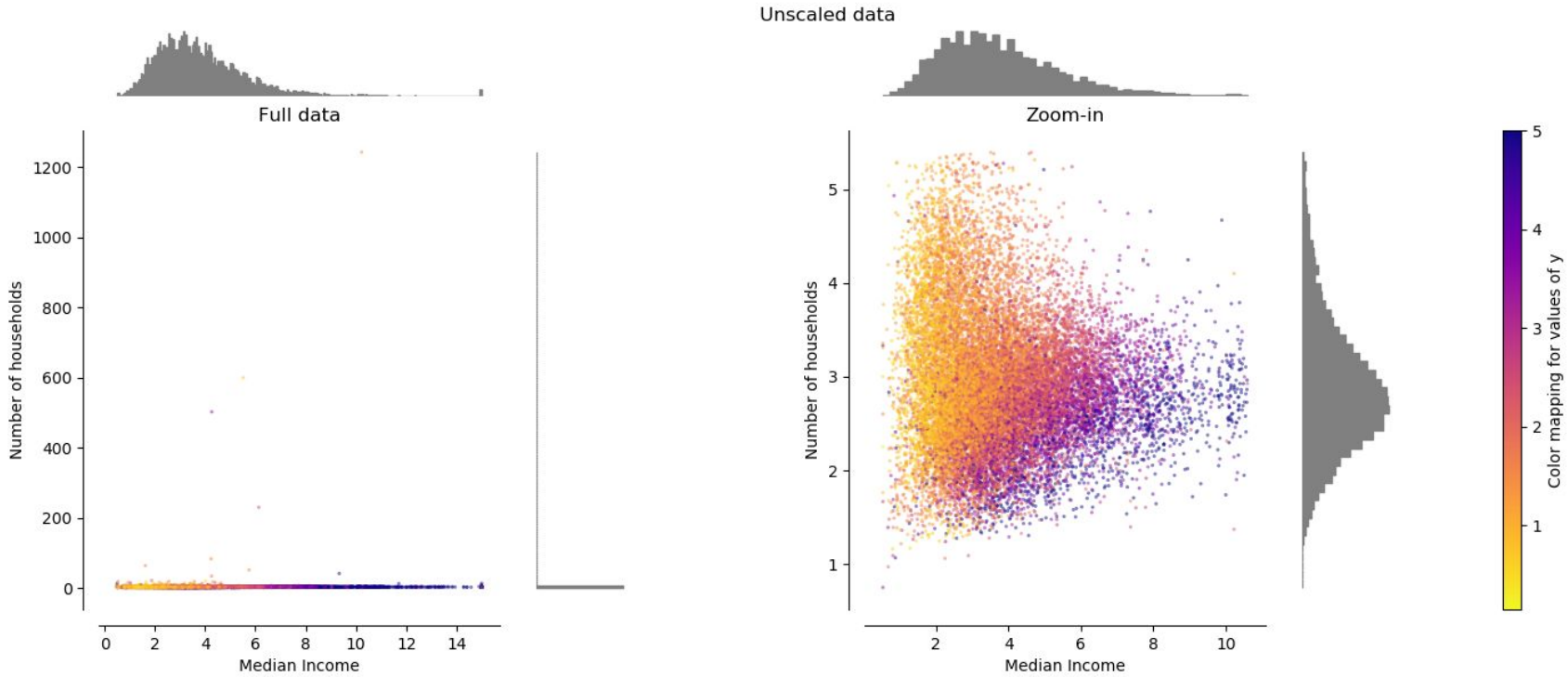
Using Weights



Modify C for each class (maybe 1/n)

MANCHESTER 1824
The University of Manchester

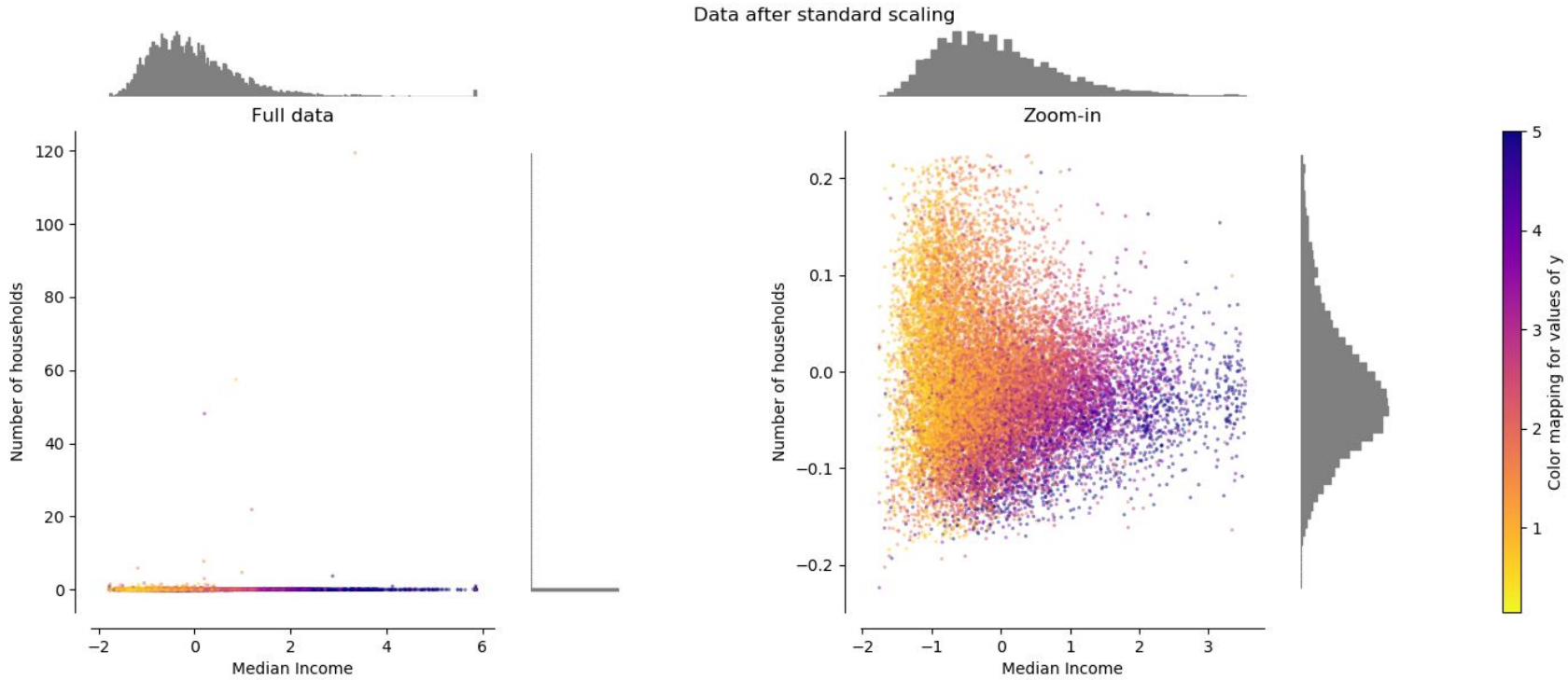
Scaling



Dr Joe Fennell | Jodrell Bank Centre for Astrophysics, University of Manchester, UK | joseph.fennell@manchester.ac.uk

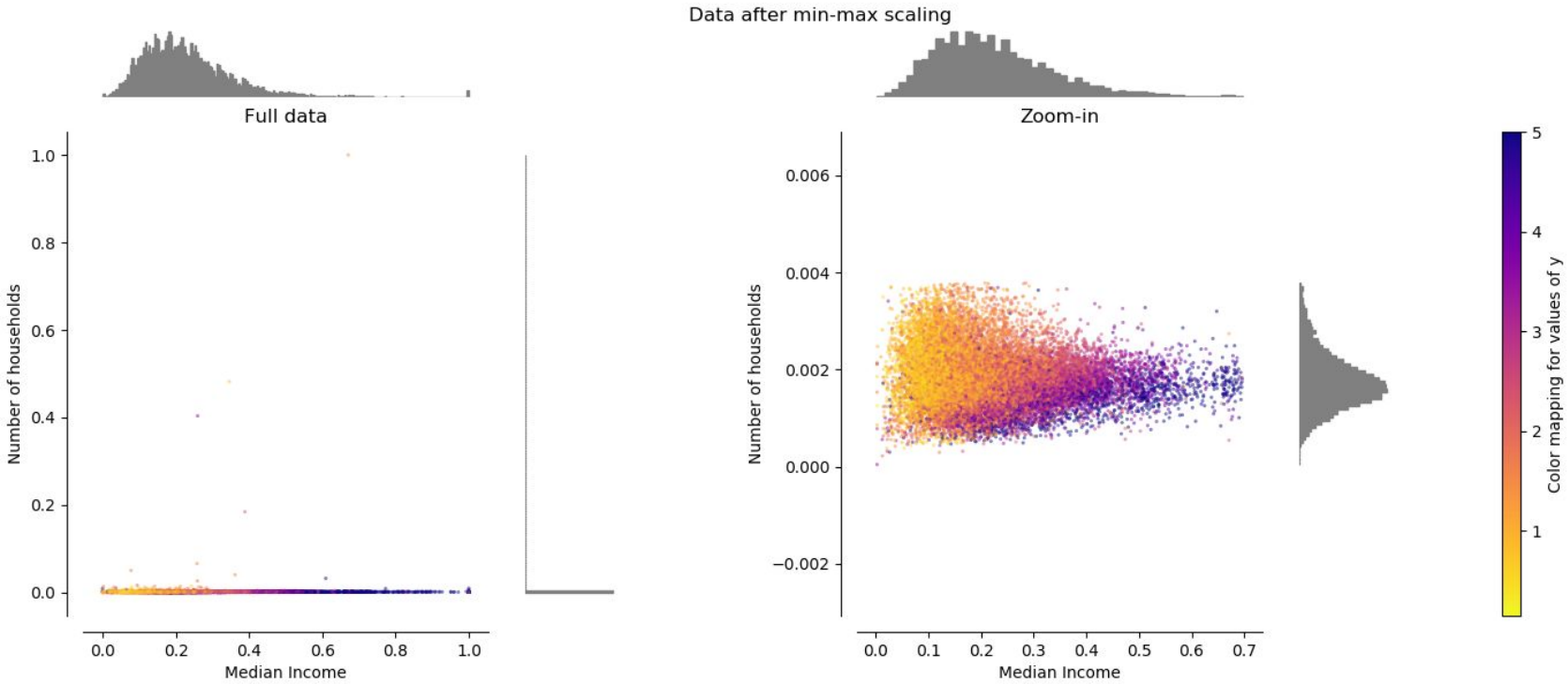
MANCHESTER 1824
The University of Manchester

Scaling: Standard



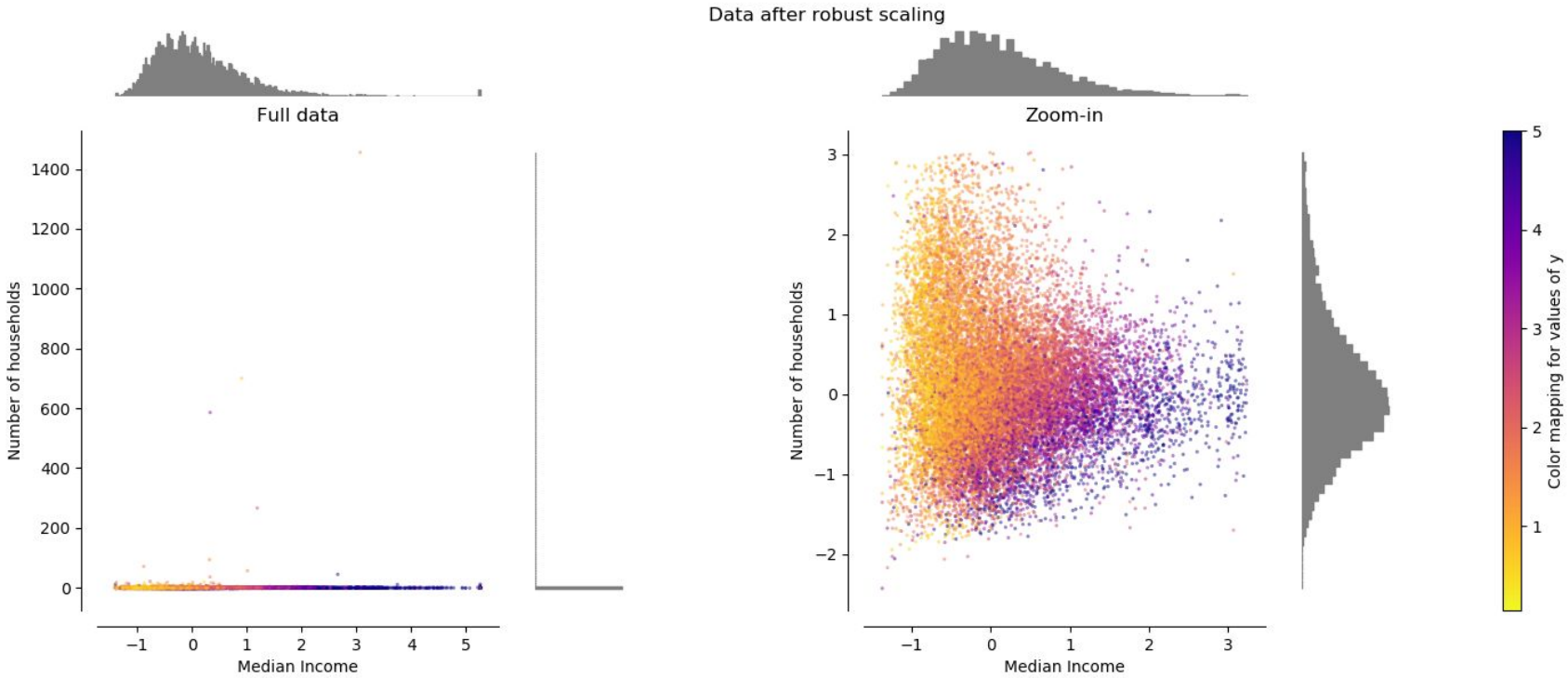
MANCHESTER 1824
The University of Manchester

Scaling: MinMax



MANCHESTER 1824
The University of Manchester

Scaling: Robust



Dr Joe Fennell | Jodrell Bank Centre for Astrophysics, University of Manchester, UK | joseph.fennell@manchester.ac.uk

MANCHESTER 1824
The University of Manchester

A real world example

Fennell *et al. Plant Methods (2018) 14:82*
https://doi.org/10.1186/s13007-018-0350-3

Plant Methods

METHODOLOGY Open Access

A method for real-time classification of insect vectors of mosaic and brown streak disease in cassava plants for future implementation within a low-cost, handheld, in-field multispectral imaging sensor

Joseph Fennell[2], Charles Veys[1], Jose Dingle[1], Joachim Nwezeobi[3], Sharon van Brunschot[3], John Colvin[3] and Bruce Grieve[1*]

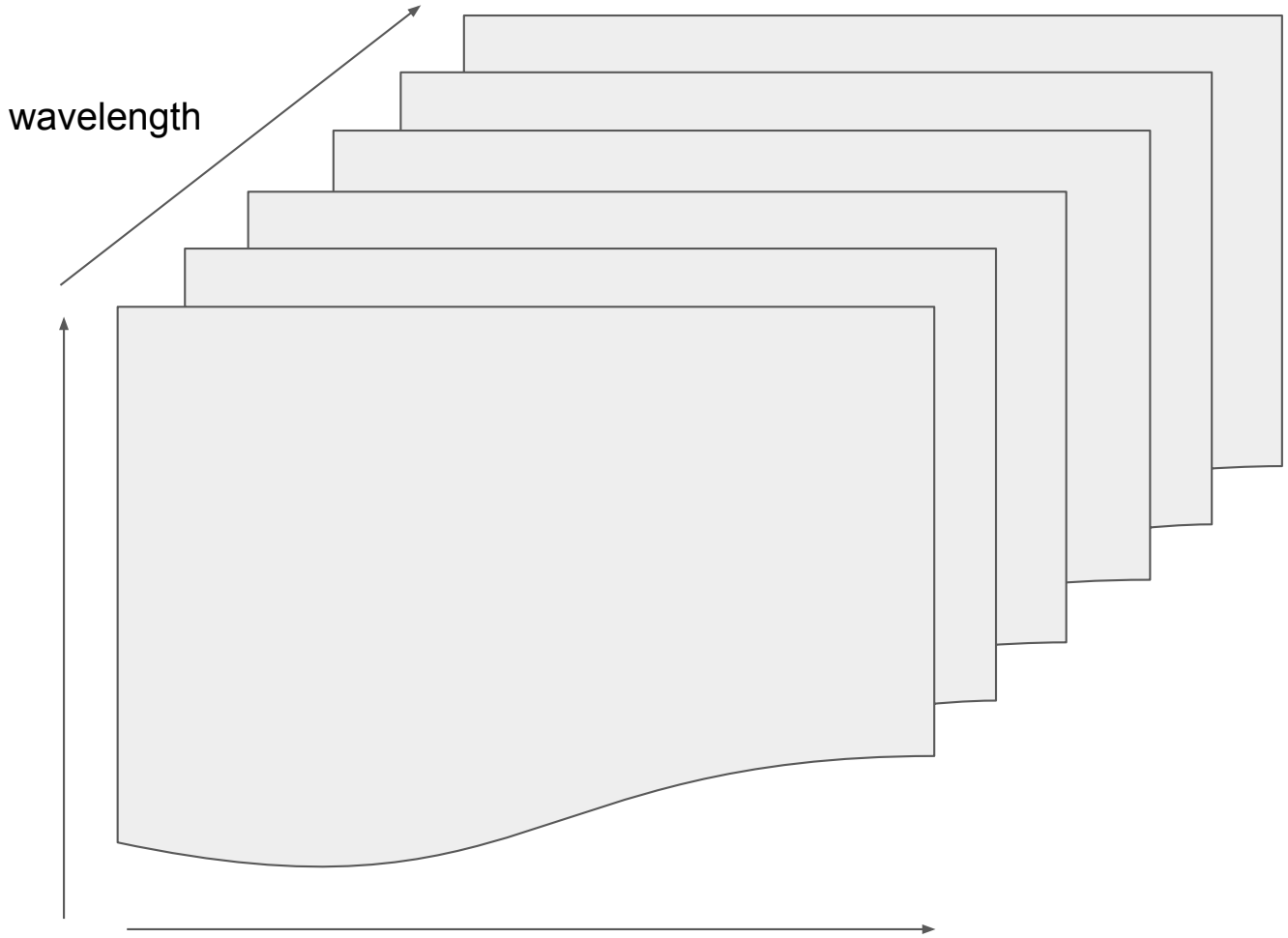
Abstract

Background: The paper introduces a multispectral imaging system and data-processing approach for the identification and discrimination of morphologically indistinguishable cryptic species of the destructive crop pest, the whitefly *Bemisia tabaci*. This investigation and the corresponding system design, was undertaken in two phases under controlled laboratory conditions. The first exploited a prototype benchtop variant of the proposed sensor system to analyse four cryptic species of whitefly reared under similar conditions. The second phase, of the methodology development, employed a commercial high-precision laboratory hyperspectral imager to recover reference data from five

Dr Joe Fennell | Jodrell Bank Centre for Astrophysics, University of Manchester, UK | joseph.fennell@manchester.ac.uk

MANCHESTER 1824
The University of Manchester

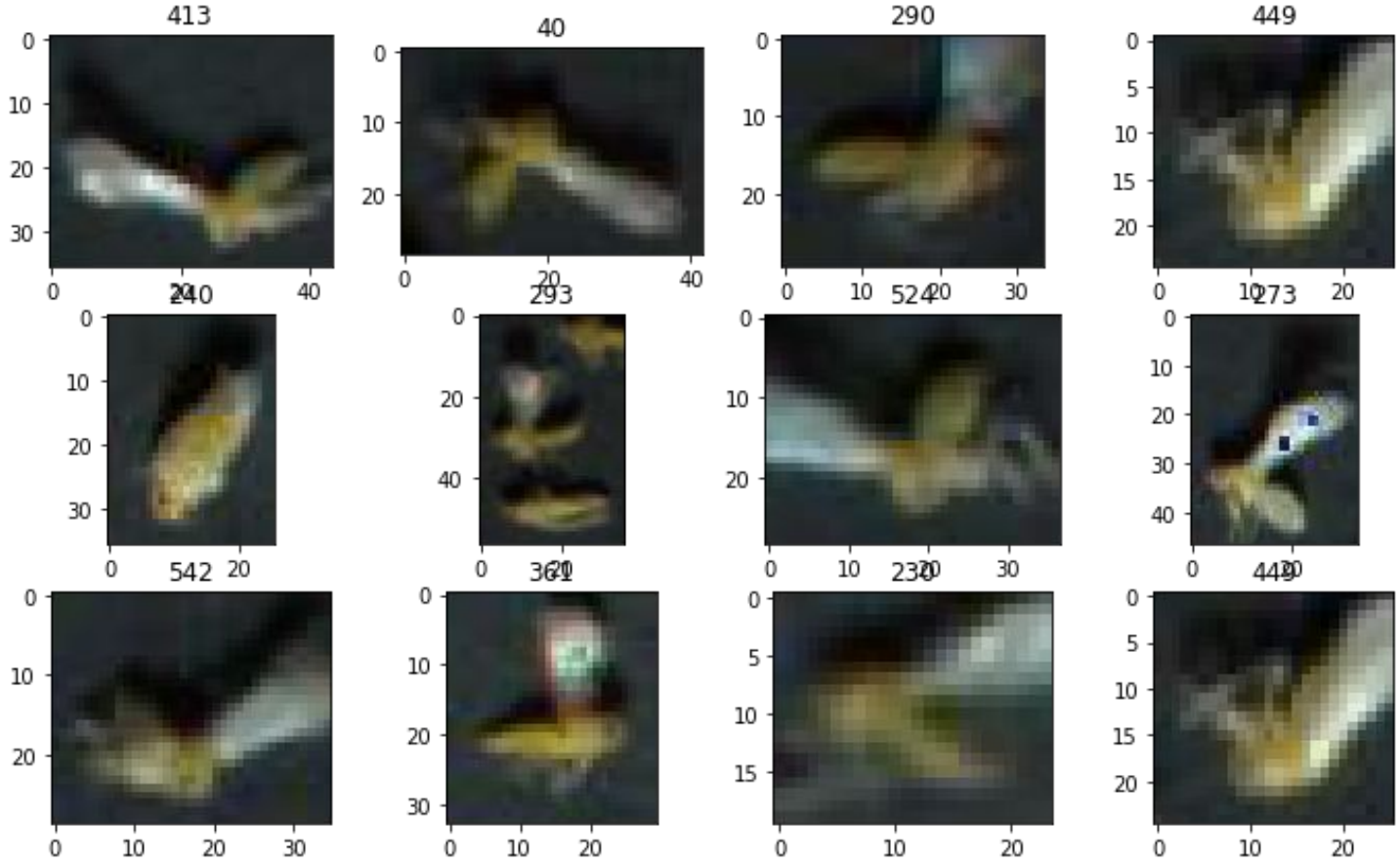
Hyperspectral Images



Dr Joe Fennell | Jodrell Bank Centre for Astrophysics, University of Manchester, UK | joseph.fennell@manchester.ac.uk

MANCHESTER 1824
The University of Manchester

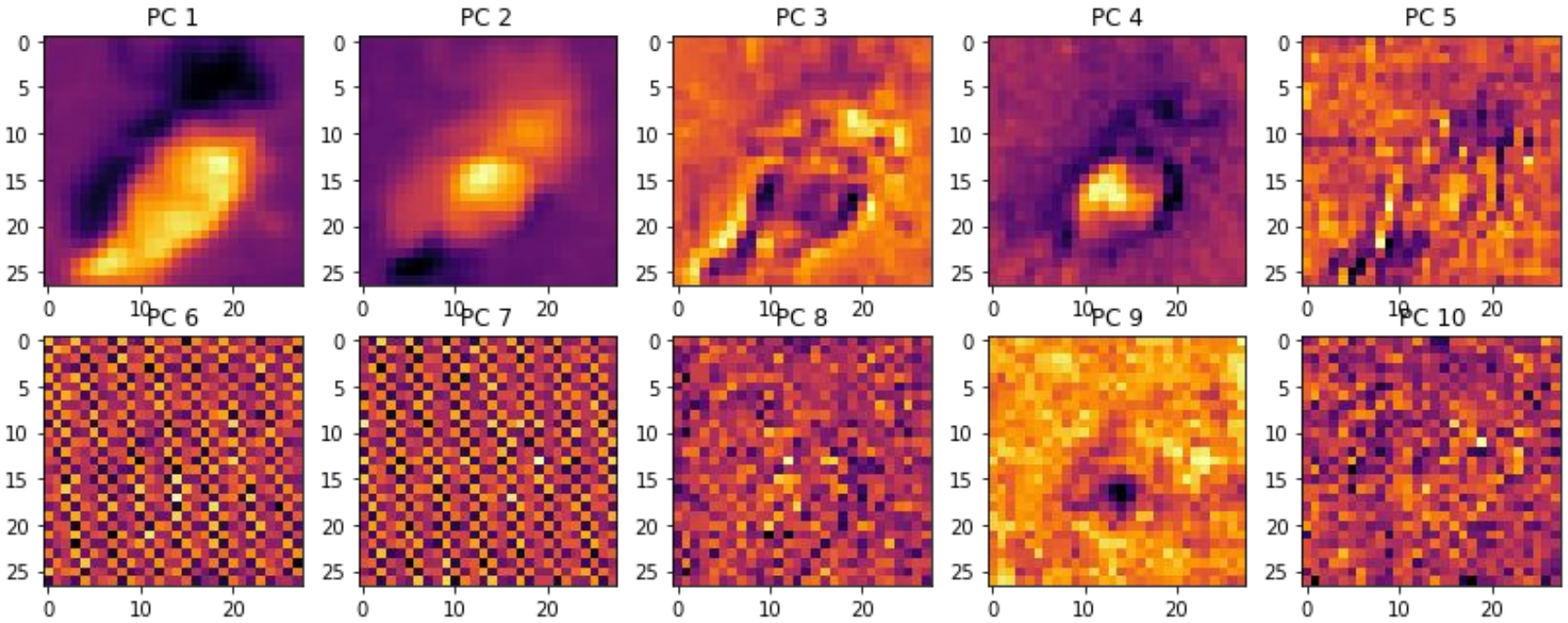
Our data



Dr Joe Fennell | Jodrell Bank Centre for Astrophysics, University of Manchester, UK | joseph.fennell@manchester.ac.uk

MANCHESTER 1824
The University of Manchester

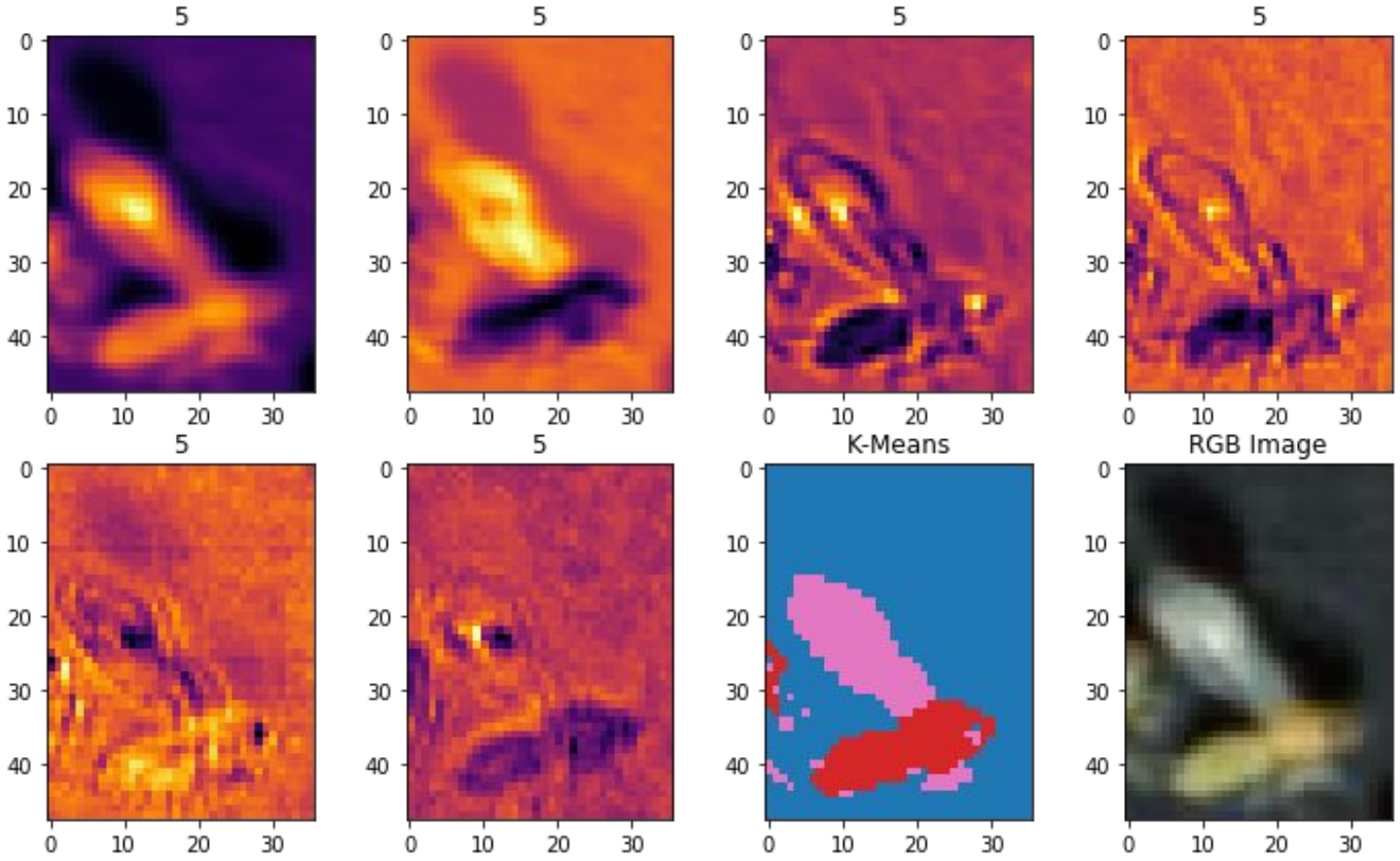
Dimensionality reduction



Dr Joe Fennell | Jodrell Bank Centre for Astrophysics, University of Manchester, UK | joseph.fennell@manchester.ac.uk

MANCHESTER 1824
The University of Manchester

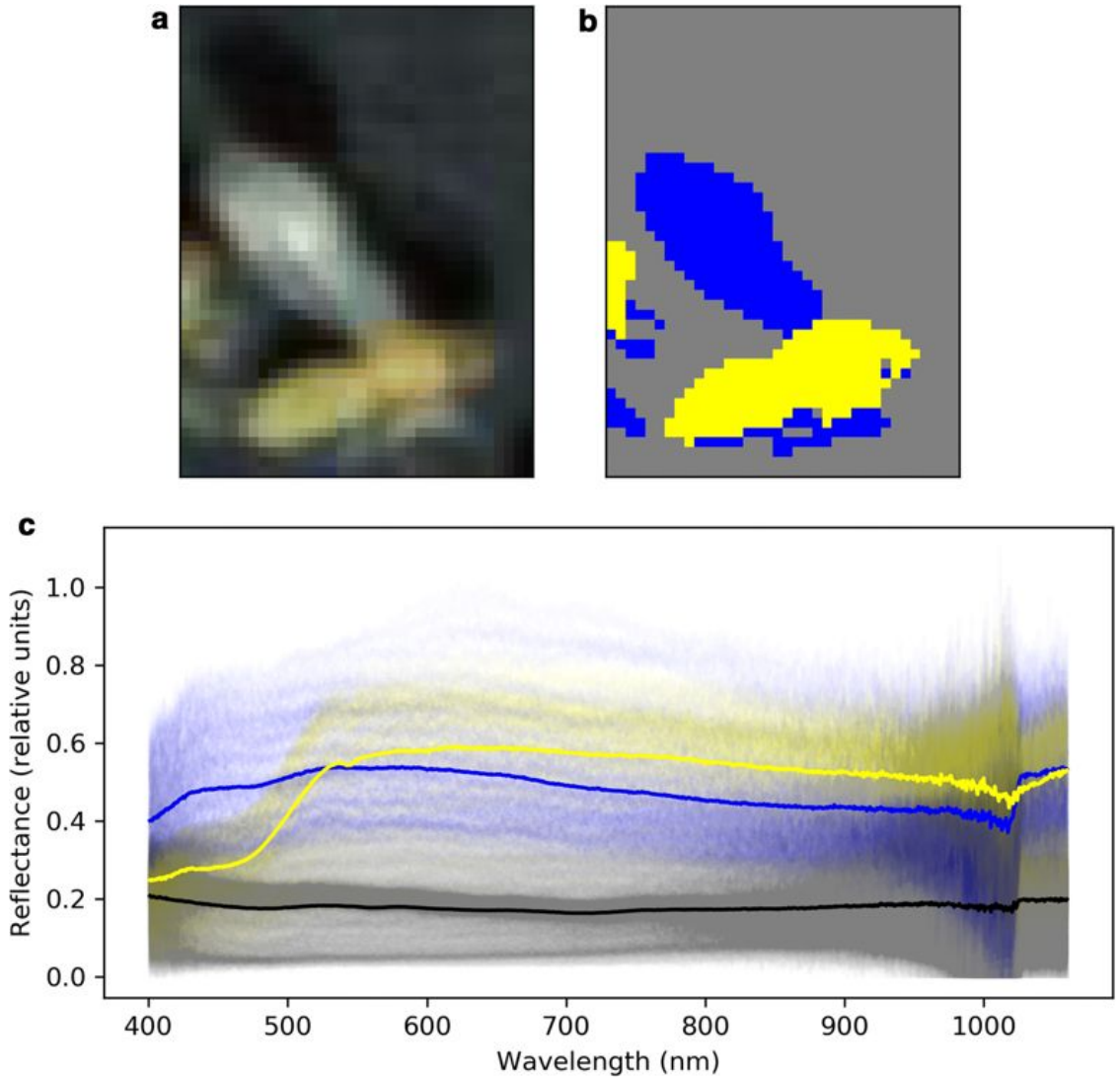
Clustering



Dr Joe Fennell | Jodrell Bank Centre for Astrophysics, University of Manchester, UK | joseph.fennell@manchester.ac.uk

MANCHESTER 1824
The University of Manchester

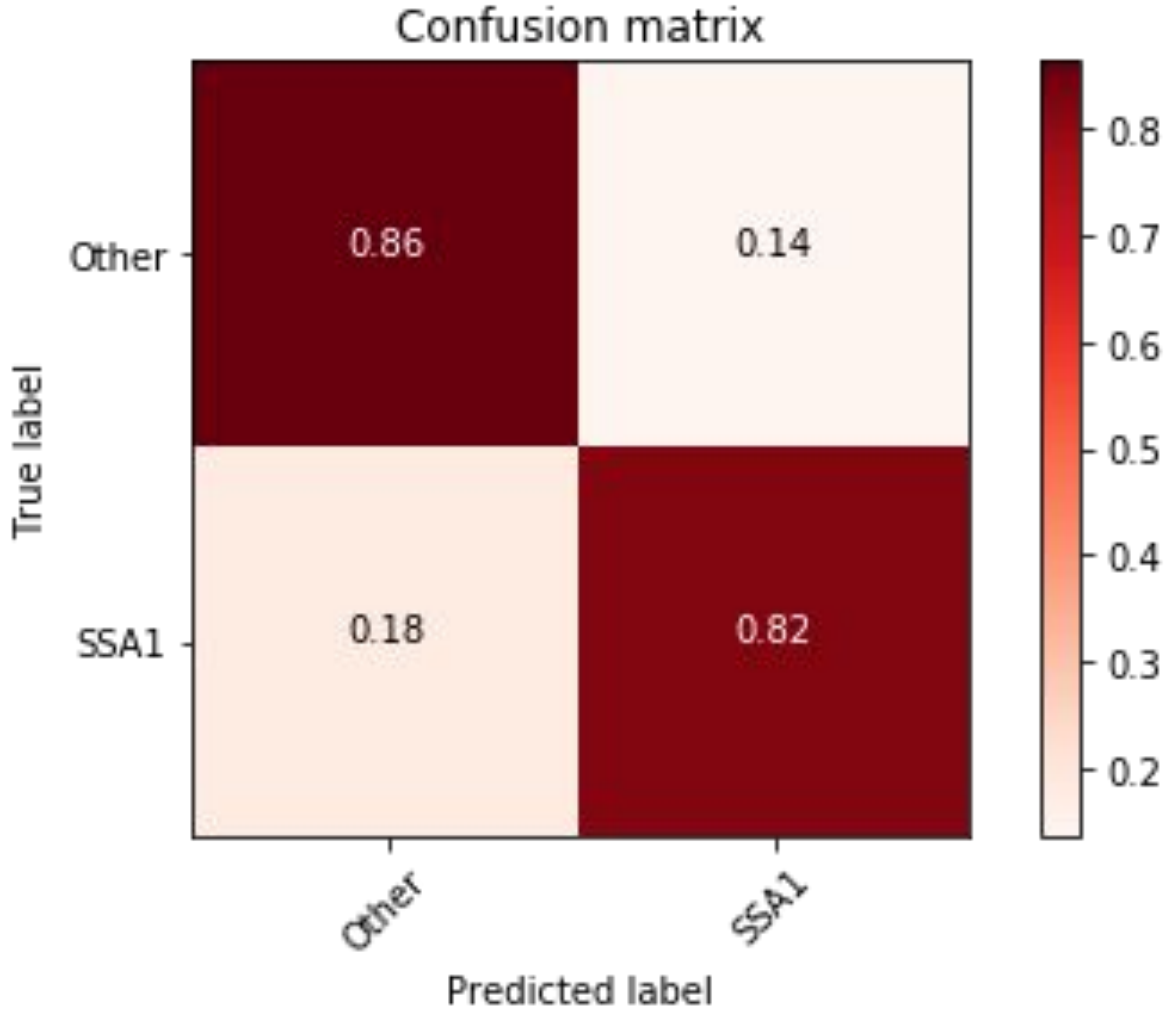
Clustering



Dr Joe Fennell | Jodrell Bank Centre for Astrophysics, University of Manchester, UK | joseph.fennell@manchester.ac.uk

MANCHESTER 1824
The University of Manchester

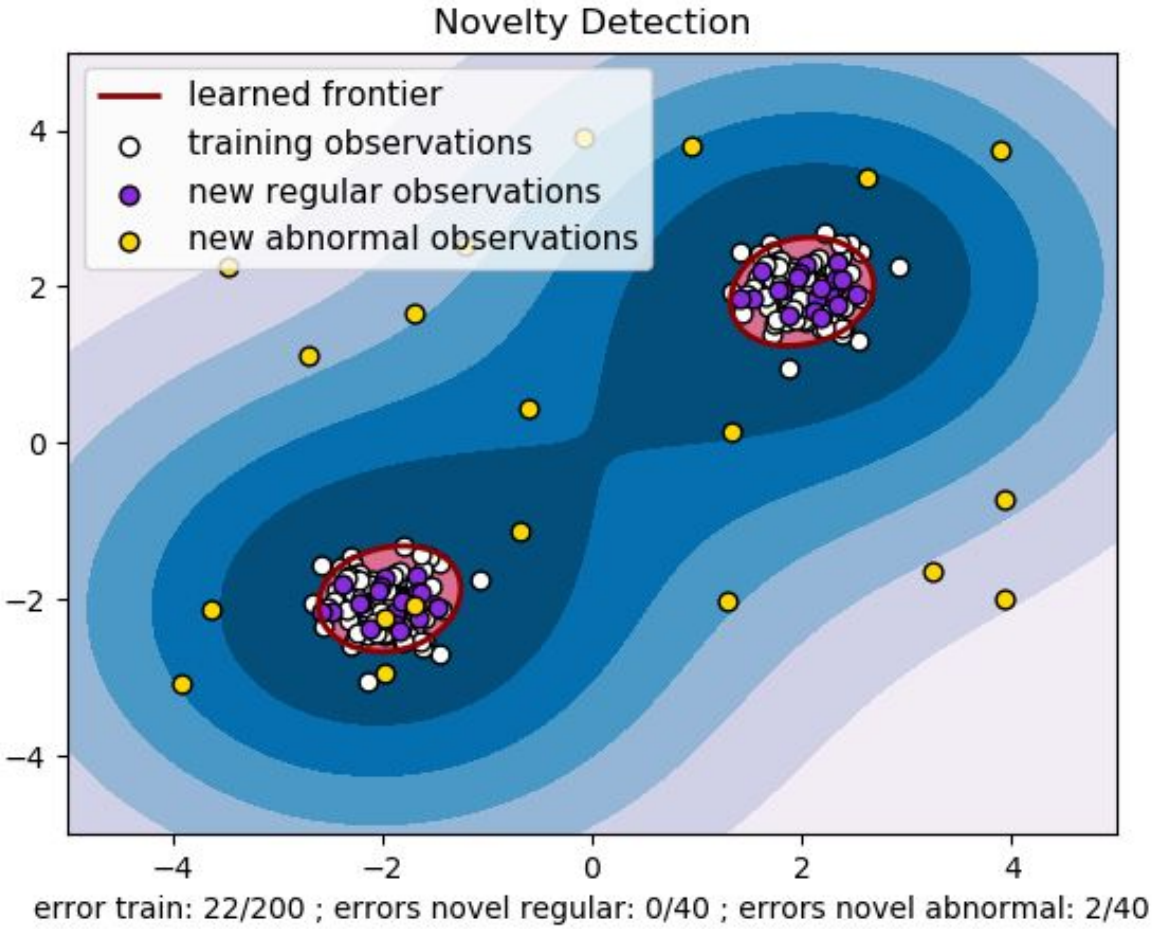
Classification



Dr Joe Fennell | Jodrell Bank Centre for Astrophysics, University of Manchester, UK | joseph.fennell@manchester.ac.uk

MANCHESTER 1824
The University of Manchester

Other types of SVM: One Class



Dr Joe Fennell | Jodrell Bank Centre for Astrophysics, University of Manchester, UK | joseph.fennell@manchester.ac.uk

MANCHESTER 1824
The University of Manchester

Multiclass SVM

- One v. The rest (OVR)
  - Fast and efficient
  - Less accurate than one v. one
- One v. One (OVO)
  - Costly algorithm to run

https://scikit-learn.org/stable/modules/svm.html