



# Data Literacy

Tutor : Ghasan Al-Falouji  
Supervisor : Prof. Dr. Roland Mandl



```
In [1]: ## CSS coloring for the dataframe tables

from IPython.core.display import HTML
#
css = open('../style/style-table.css').read() + open('../style/style-notebook.css').read()
HTML('<style>{}/</style>'.format(css))
```

Out[1]:

## Introduction to Classification

In [ ]:

## Logistic Regression

logistic-regression classifiers decision boundaries on the first two dimensions (sepal length and width) of the [iris dataset](#). The datapoints are colored according to their labels.

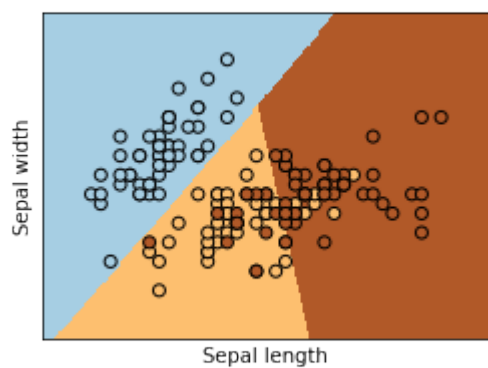
```
In [1]: import numpy as np
import matplotlib.pyplot as plt
from sklearn.linear_model import LogisticRegression
from sklearn import datasets

%matplotlib inline
```

```
In [2]: # import some data to play with
iris = datasets.load_iris()
X = iris.data[:, :2] # we only take the first two features.
Y = iris.target

##### CODE HERE #####
```

Out[2]: ([], <a list of 0 Text yticklabel objects>)



## k-NN Classification

Using the previously imported iris data, we will apply now kNN

```
In [10]: import numpy as np
import matplotlib.pyplot as plt
from matplotlib.colors import ListedColormap
from sklearn import neighbors, datasets
```

```
In [12]: # import some data to play with
iris = datasets.load_iris()

# we only take the first two features. We could avoid this ugly
# slicing by using a two-dim dataset
X = iris.data[:, :2]
y = iris.target

##### CODE HERE #####
```

