Python Libraries required for Dr Rob Collins' Machine Learning Course

Dr Rob Collins

Version 11, 23rd August 2023

(c) Donox Ltd 2021

This notebook lists the libraries required for each session. If you can run every code block within each session without an error showing, then you have the required libraries installed within your environment.

Note : I use this 'Required Libraries' notebook for several of my courses. Depending on the course you are completing you may not be asked to complete all of the sessions below. So please check which sessions you will actually be completing before spending excessive time installing libraries in your local environment.

Note that libraries are only listed once in this list. Thus, for example, if the libraries used in sessino 1 are then re-used in the Session 4 tutorial I do not re-list them. Many of the libraries that are introduced early in the course are re-used later in the course. You should therefore check that you have access to all of these libraries.

Session 0 : Python Re-fresh

```python
In [ ]:  import math
```

```python
In [ ]:  import random
```

```python
In [ ]:  import datetime
```

Session 0a : Matricies and Linear Algebra

This is an optional module for students who wish to further their understanding of Linear Algebra and its relationship to Machine Learning

```python
In [ ]:  import sympy as smp          # sympy is an elegant library for symbolic linear algebra.
         import numpy as np           # Numpy is also included below and used extensively in the course
```

Session 1 : Data Cleansing

```python
In [ ]:  import pandas                 # Allows easy creation and management
                                       # of dataframes (data tables).
                                       # See https://pandas.pydata.org/
```

```python
In [ ]:  import numpy                  # Powerful functions for numerical
                                       # computation. See https://numpy.org/
```

```python
In [ ]:  import seaborn as sns         # Statistical data visualisation.
                                       # See http://seaborn.pydata.org
```

```python
In [ ]:  import matplotlib             # Data plotting library.
                                       # See https://matplotlib.org/
```

```python
In [ ]:  import missingno     # Library to visualise missing data.
                              # See https://github.com/ResidentMario/missingno
               # Install using:  conda install -c conda-forge missingno
```

Session 2 : End-to-end example of supervised learning

```python
In [ ]:  import sklearn        # A powerful and extensive Machine
                               # Learning library.
                               # See https://scikit-learn.org/stable/index.html
```

```python
In [ ]:  import pickle         # Enables saving and loading of models to binary files
                               # Useful when it takes a long time to build a model .. which can
                               # then be saved and re-used at a later date.
```

Session 3a : Clustering

```python
In [ ]:  import mpl_toolkits          # A simple 3D plotting library
                                      # extension to matplotlib.
               # See https://matplotlib.org/2.2.2/mpl_toolkits/index.html

         # The following is optional and only used at the end of the worksheet within the
         from sklearn.cluster import DBSCAN
         from sklearn.cluster import OPTICS, cluster_optics_dbscan

         import plotly.graph_objects as go
```

Session 3b : Dimensionality reduction

```python
In [ ]:  import random as rand        # Enables generation of random numbers
                                      # with a variety of distributions
```

Session 4 : Gradient descent

All required libraries are allready included in the above list

Session 5 : Polynomial regression and ROC

```python
In [ ]:  from sklearn.linear_model import LinearRegression
         from sklearn.metrics import mean_squared_error, r2_score
```

Session 6 : Trees and Ensemble

Within this workshop we will be generating pictures of decision trees. You will have to instal the 'Graphviz' tool onto your computer to generate those images:

https://graphviz.org/

Other required libraries are as follows:

```python
In [ ]:  import numpy as np
         import pandas as pd
         from pandas.plotting import scatter_matrix
         import matplotlib.pyplot as plt
         import seaborn as sns
         from sklearn.model_selection import train_test_split
         from sklearn.tree import DecisionTreeClassifier
         from sklearn.ensemble import RandomForestClassifier
         from sklearn.metrics import confusion_matrix
         from sklearn.metrics import plot_confusion_matrix
         from sklearn.model_selection import GridSearchCV
```

Session 7 : Gaussian Mixture Models

This session uses libraries that have already been introduced above.

```python
In [ ]:  import numpy as np
         import pandas as pd
         import matplotlib.pyplot as plt
```

Session 8 : Natural Language Processing

8.1 Wordcloud

The 'wordcloud' library was not part of my Anaconda environment. I thus had to instal this libary using the command:

conda install -c conda-forge wordcloud

Details of this can be found here: https://anaconda.org/conda-forge/wordcloud and here: https://amueller.github.io/word_cloud/generated/wordcloud.WordCloud.html

```python
In [ ]:  from wordcloud import WordCloud, STOPWORDS
         import matplotlib.pyplot as plt
```

8.2 Natural Language Toolkit (NLTK)

This workshop makes extensive use of the 'Natural Language Toolkit (NLTK). You can read more about NLTK here: https://www.nltk.org/

If NLTK is not part of the basic instal of your Anaconda, then you can find out how to instal it here: https://anaconda.org/anaconda/nltk

```python
In [ ]:  import nltk
         from nltk.corpus import stopwords
         from nltk.corpus.reader import tagged
```

8.3 Vader Sentiment Analayzer

https://anaconda.org/conda-forge/vadersentiment

conda install -c conda-forge vadersentiment

```python
In [ ]:  from vaderSentiment.vaderSentiment import SentimentIntensityAnalyzer
```

8.4 Other libraries used in the Natural Language Workshop

```python
In [ ]:  from collections import OrderedDict
         import csv
         import pandas as pd
         import numpy as np
         from sklearn.feature_extraction.text import CountVectorizer
```

Session 9 : Deep Learning

9.1 Keras

Since keras is not part of the base instal within Anaconda, I had to first update my environment to instal it. Thus, if you are using Anaconda you may need to open a terminal window and instal keras using the command:

conda install -c conda-forge keras

Which is described here:https://anaconda.org/conda-forge/keras

I have a high-power gpu on my computer which makes processing much faster and thus I used this alternate command:

conda install -c anaconda keras-gpu

Which is described at https://anaconda.org/anaconda/keras-gpu

This website provides a useful tutorial on how to instal tensorflow and keras: https://www.innovationmerge.com/2020/12/21/Install-TensorFlow-and-Keras-on-GPU-using-Anaconda-Navigator/

```python
In [ ]:  from sklearn.datasets import make_blobs
         from sklearn.linear_model import Perceptron
         from sklearn.model_selection import train_test_split
         from sklearn.metrics import accuracy_score
         import numpy as np
         import matplotlib.pyplot as plt

         import tensorflow.keras
         from tensorflow.keras.models import Sequential
         from tensorflow.keras.layers import Conv2D,MaxPooling2D,Dense,Flatten,Dropout
         from tensorflow.keras.optimizers import Adam
```

Session 10 : Reinforcement Learning

AI Gym can be installed in the Anaconda Envionment using the command:

```
conda install -c conda-forge gym
```

Required libraries are:

```python
In [ ]:  import gym
         from IPython.display import import clear_output
```

```python
In [ ]:
```