SPWLA 2021: Machine Learning & AI Workshop

Instructors: Lalitha Venkataramanan (Schlumberger), Chicheng Xu (Aramco), Andy McDonald (Lloyd's Register), Vikas Jain (Schlumberger)

Introduction

This workshop will focus on the applications of Artificial Intelligence (AI) and Machine Learning (ML) to the upstream O&G industry. Consisting of two half-days, the workshop will introduce machine learning, lay out sample workflows and steps for ML applications and summarize some of the used cases in the industry.

Course Material

Link to the workshop Repository:
https://github.com/SPWLA-ORG/spwla2021_ml_workshop

Link to Interactive Binder Notebooks:
https://mybinder.org/v2/gh/SPWLA-ORG/spwla2021_ml_workshop/HEAD

Machine Learning Types

Supervised Learning
Most common machine learning task. It is designed to learn by example using input data that has been paired with the correct outputs. After the model has been trained it can be used to predict an output.

Can be split up into regression, where models are used to predict continuous numerical output, and classification, where models are used to predict a discrete output category.

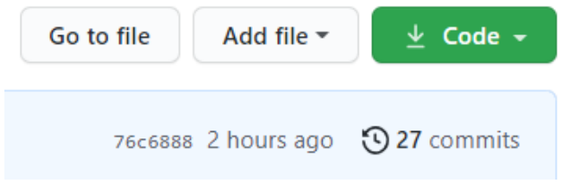
Unsupervised Learning
Used to identify underlying patterns within the data without the need for labelled data. It can be used for initial exploratory data analysis, dimensionality reduction and classification

Reinforcement Learning
Goal orientated algorithms that learn to make decisions in order to achieve complex objectives based on interactions with its environment.

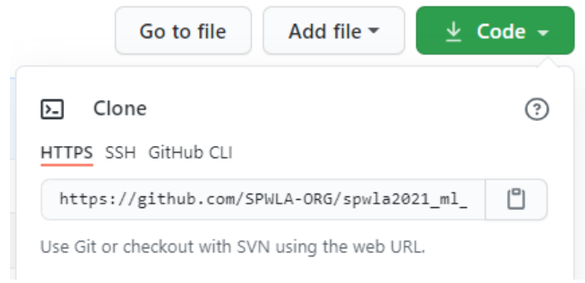
Downloading Course Material

Cloning the Repository

1. Navigate to the GitHub Repository
2. In the top right, click on the green Code button

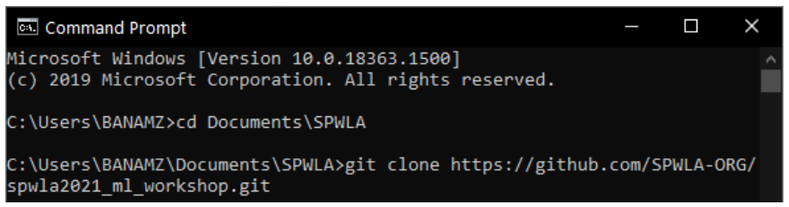


3. Click on the copy button to copy the repository address



4. Open a terminal window (MacOS) or command prompt window (Windows)

5. Navigate to a directory where you want the repository to be downloaded to.
6. Type in *'git clone https://github.com/SPWLA-ORG/spwla2021_ml_workshop.git'*



7. Once the data has downloaded you will be able to access it from a Jupyter Notebook or from Jupyter-Labs

Data

The data used for this workshop comes from the Volve dataset.

In 2018, Equinor released the entire contents of the Volve Field to the public domain to foster research and learning. Data includes: Well Logs; Petrophysical interpretations; Reports; Core Measurements; Seismic data; etc.

The Volve Field is located some 200 km west of Stavanger in the Norwegian Sector of the North Sea. Hydrocarbons were discovered within the Jurassic aged Hugin Formation in 1993. Oil production began in 2008 and lasted for 8 years (twice as long as planned) until 2016, when production ceased. In total 63 MMBO were produced over the field's lifetime and reached a plateau of 56,000 B/D.

Details for the Volve Field and the entire dataset can be found here. The full license agreement can be found here.

Curve Name	Units	Description
MD	m	Measured Depth
BS	in	Bitsize
CALI	in	Caliper
DT	us/ft	Acoustic Compressional Slowness
DTS	us/ft	Acoustic Shear Slowness
GR	api	Gamma Ray
NPHI	dec	Neutron Porosity
RACEHM	ohm.m	Resistivity (High Freq. Atten)
RACELM	ohm.m	Resistivity (Low Freq. Atten)
RHOB	g/cc	Bulk Density
RPCEHM	ohm.m	Resistivity (High Freq. Phase)
RPCELM	ohm.m	Resistivity (Low Freq. Phase)
PHIF	dec	Final Porosity
SW	dec	Water Saturation
VSH	dec	Shale Volume

Andy McDonald • andrew.mcdonald@lr.org • Updated: 2021-05

SPWLA 2021: Machine Learning & AI Workshop

Instructors: Lalitha Venkataramanan (Schlumberger), Chicheng Xu (Aramco), Andy McDonald (Lloyd's Register), Vikas Jain (Schlumberger)

Libraries Used in Workshop

Data Storage & Manipulation

pandas NumPy

Data Visualisation

matplotlib seaborn

Machine Learning

TensorFlow K scikit learn

Working With Libraries

Installing Libraries

```
pip install numpy
pip install keras
pip install matplotlib
pip install seaborn
```

Importing Libraries

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from math import pi
```

Installing From requirements.txt

Open Command Prompt

CD to folder where reqtuirements.txt exists

Type:

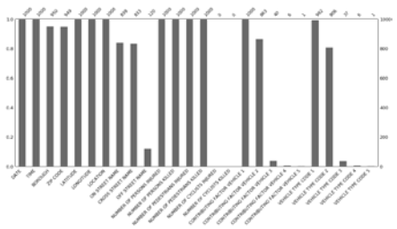
pip install –r requirements.txt

pandas

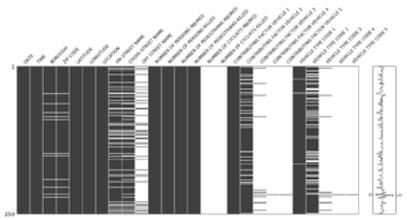
df = pd.read_csv('file.csv')	Load in a csv file
df.describe()	View key stats about the data
df.info()	View a summary of the dataframe
df.head(n)	View the first n rows of the dataframe
df.tail(n)	View the last n rows of the dataframe
df['GR']	Access a column by name
df.iloc[:,1]	Access a column by index position
df.iloc[1,:]	Access a row by index position
df.dropna()	Drop all null values
df.drop('GR', axis=1)	Drop a column

missingno

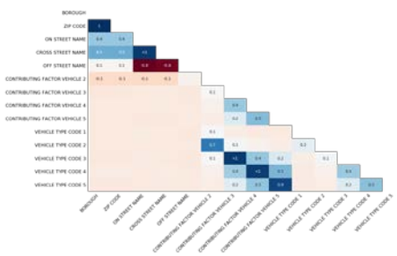
msno.bar(df)



msno.matrix(df)



msno.heatmap(df)



msno.bar(df)

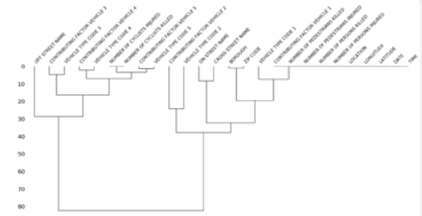


Image from https://github.com/ResidentMario/missingno

matplotlib

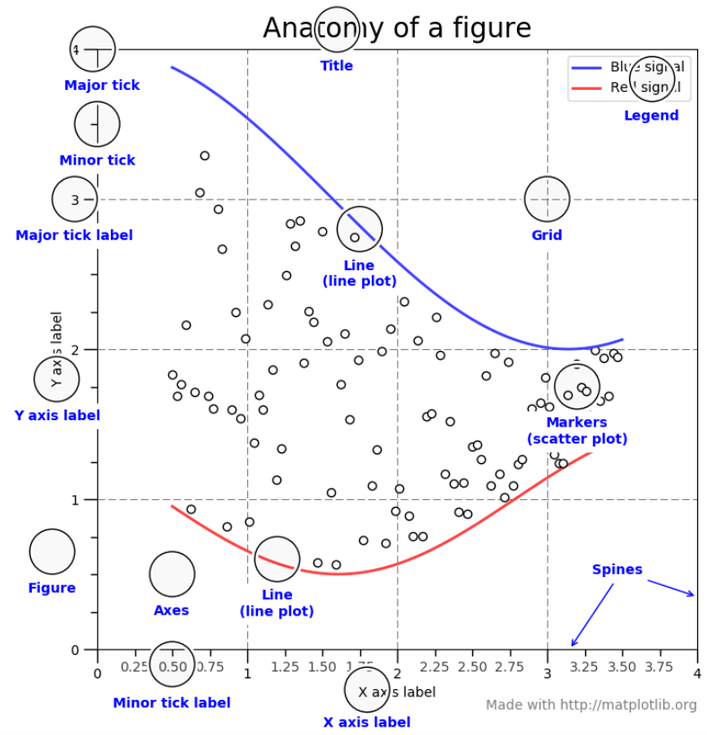


Image from https://matplotlib.org/stable/gallery/showcase/anatomy.html

Creating plots

plt.plot(x, y)	Create a line plot
plt.scatter(x,y)	Create a scatter plot
plt.boxplot(x)	Create a boxplot
plt.bar(x)	Create a bar plot
plt.hist(x)	Create a histogram
plt.violin(dataset)	Create a violin plot

Andy McDonald • andrew.mcdonald@lr.org • Updated: 2021-05