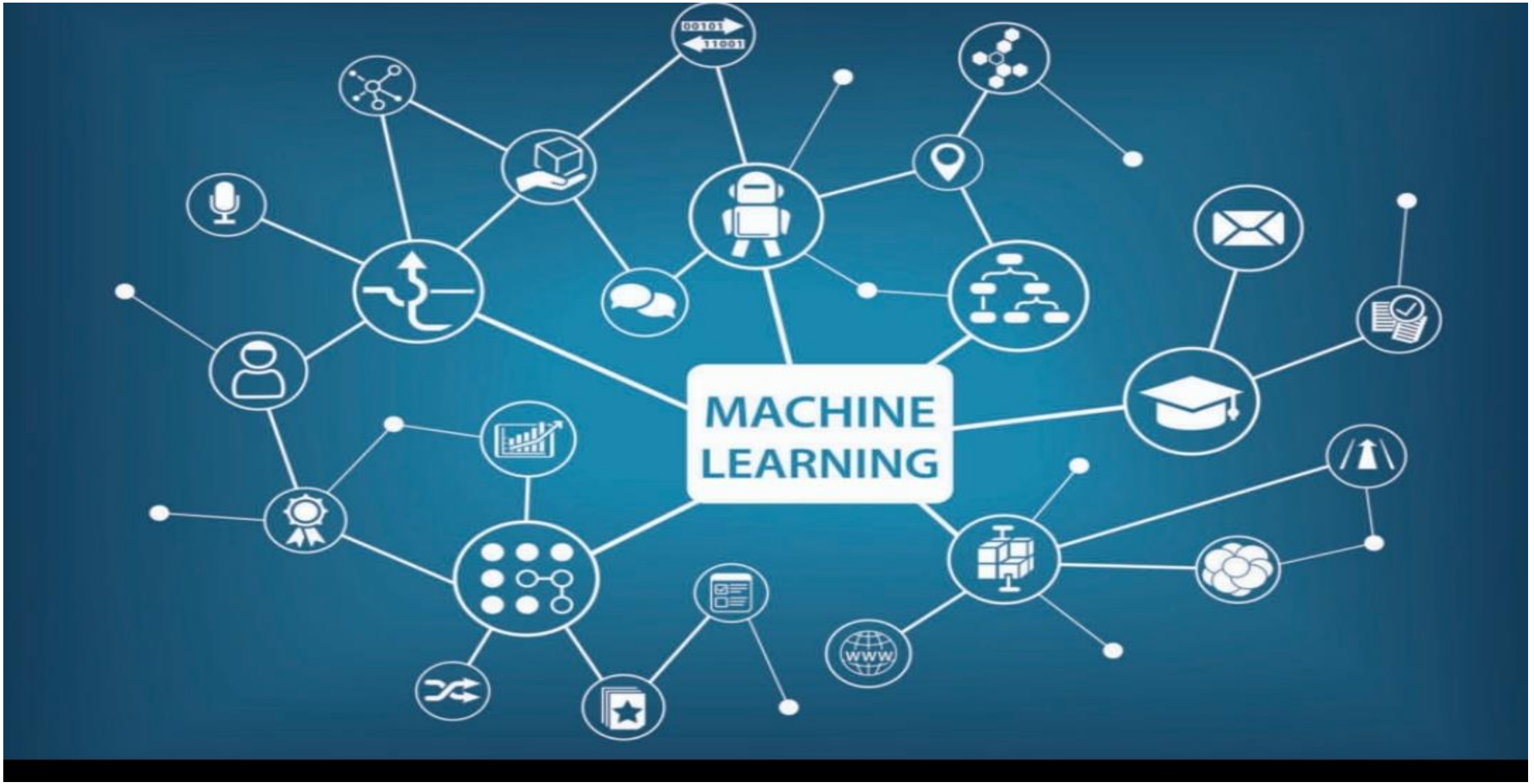
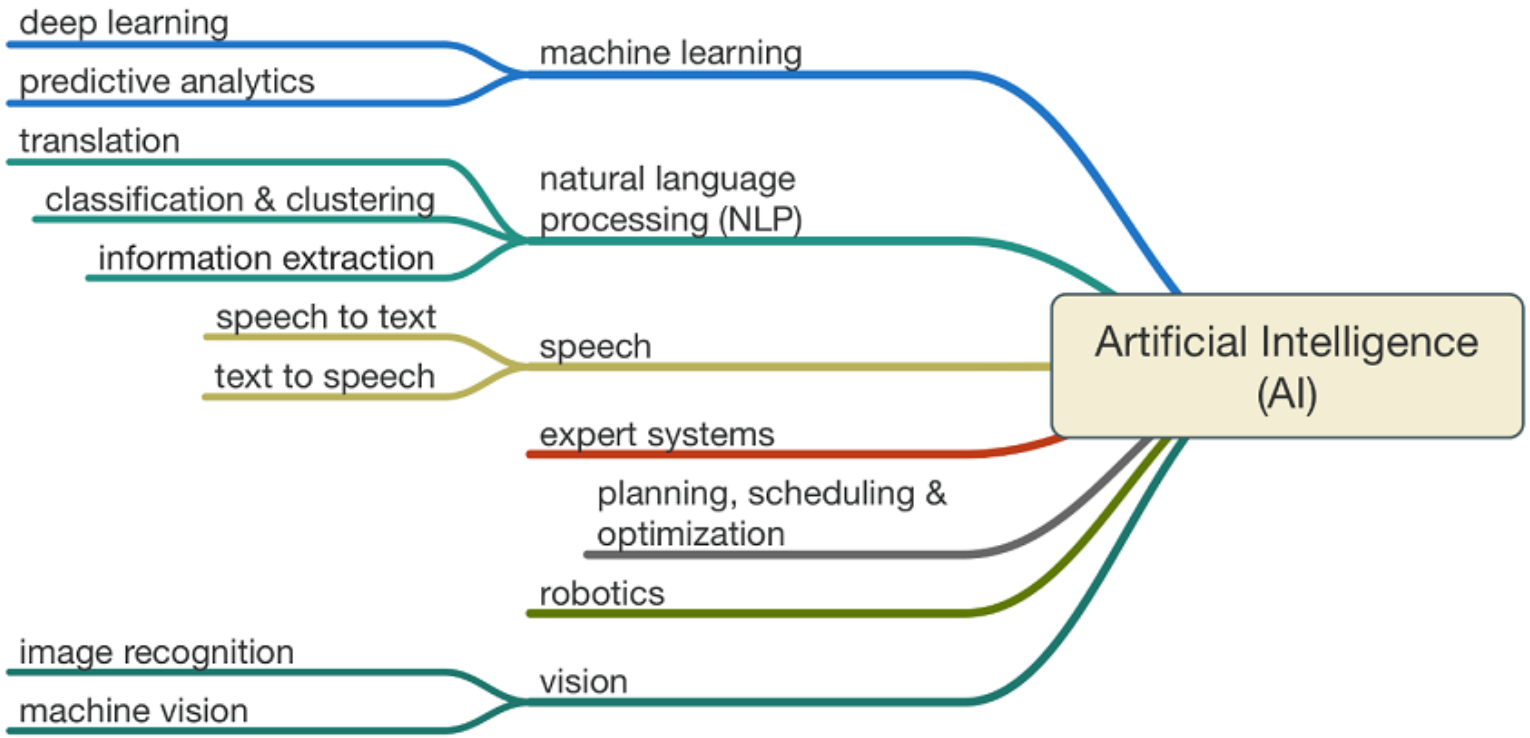
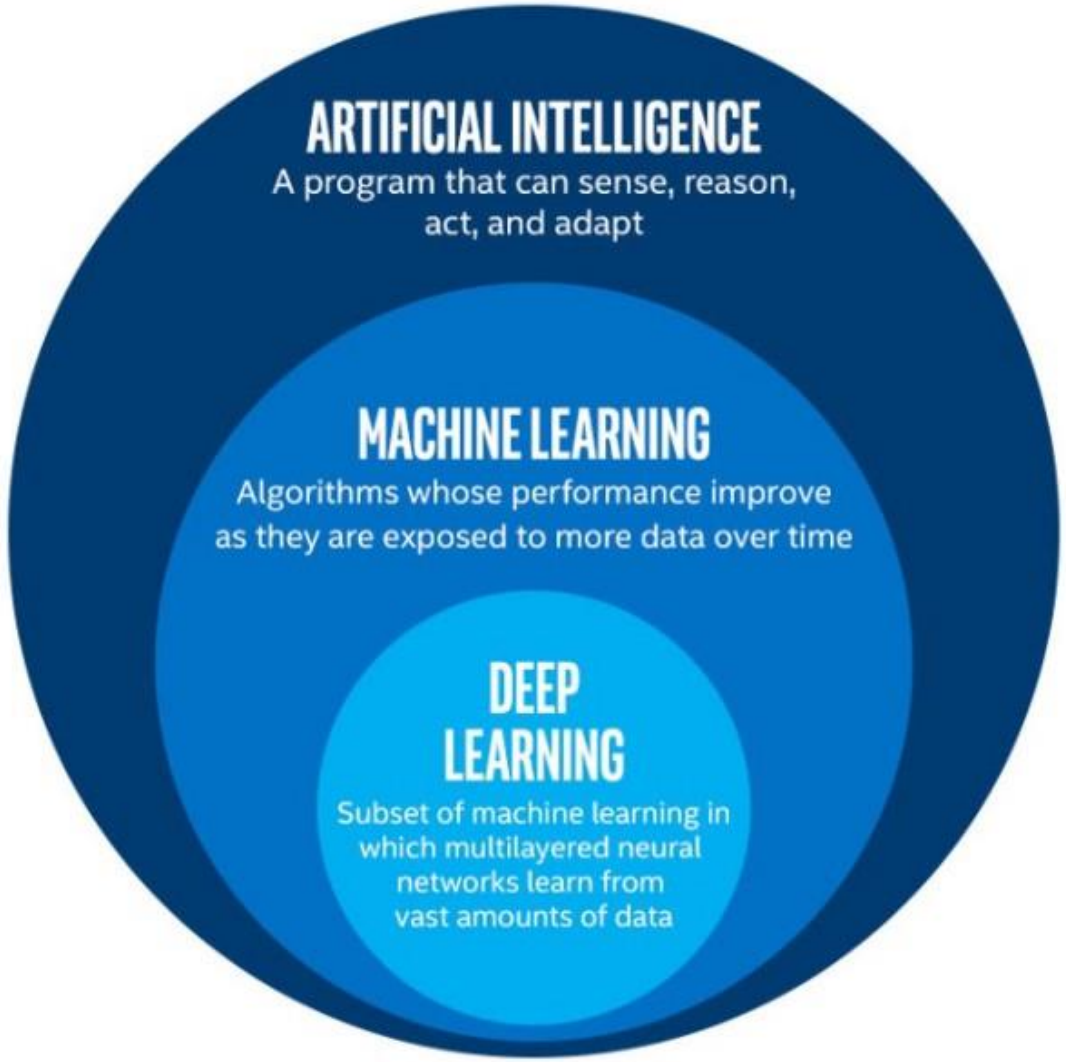

MACHINE
LEARNING
WWW

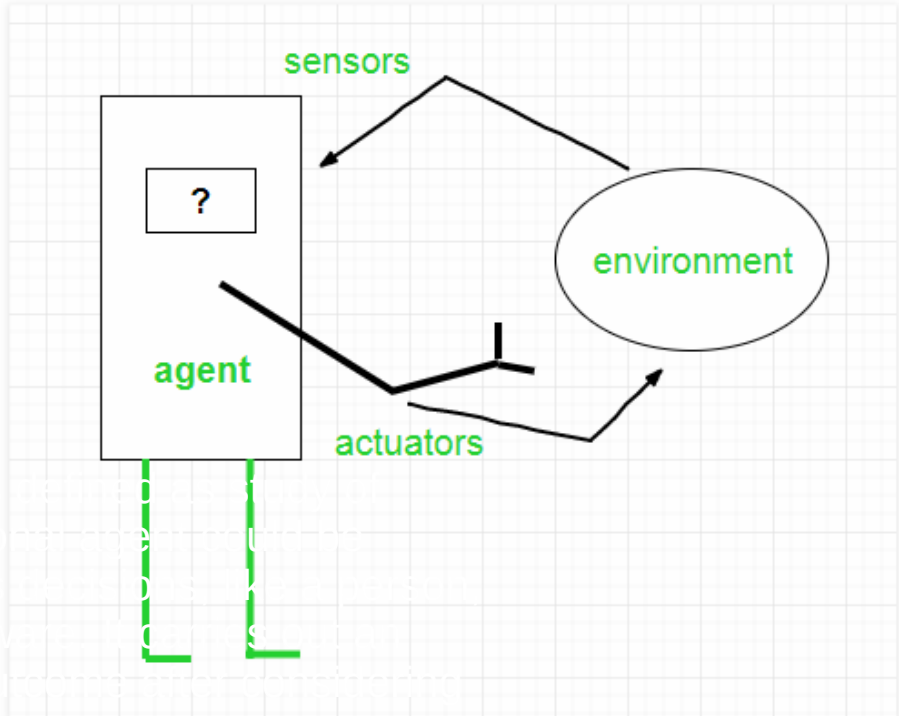
Artificial intelligence (AI) is an area of computer science that emphasizes the creation of intelligent machines that work and react like humans.



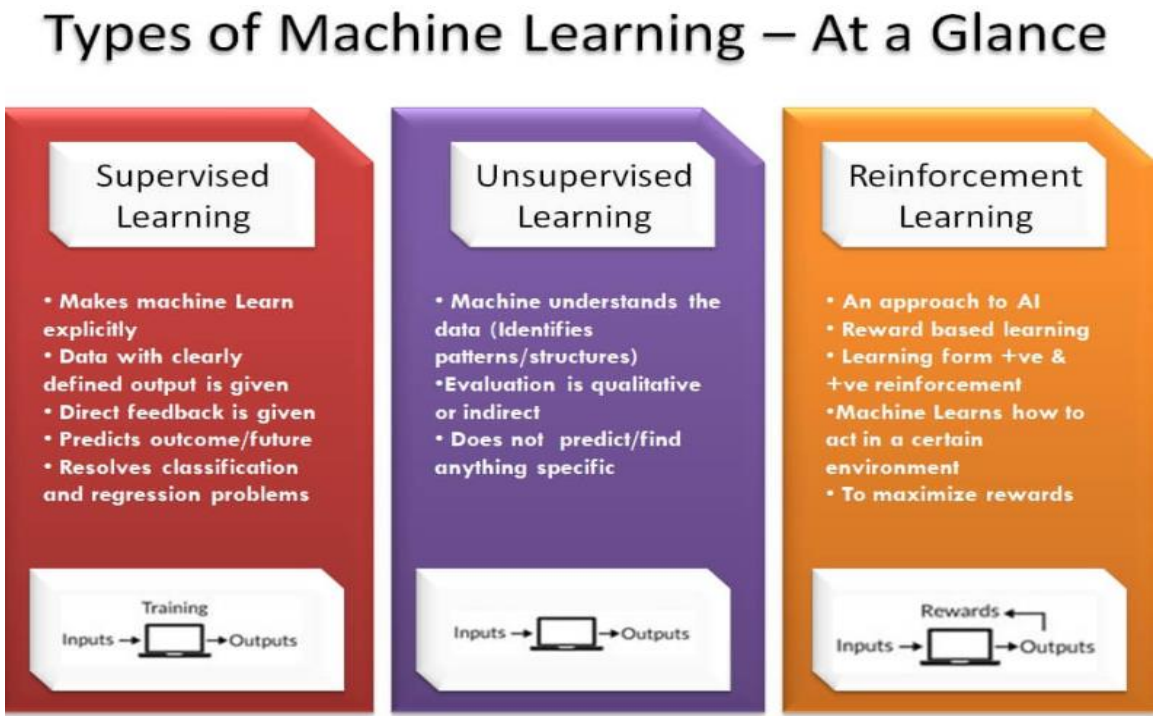
Examples

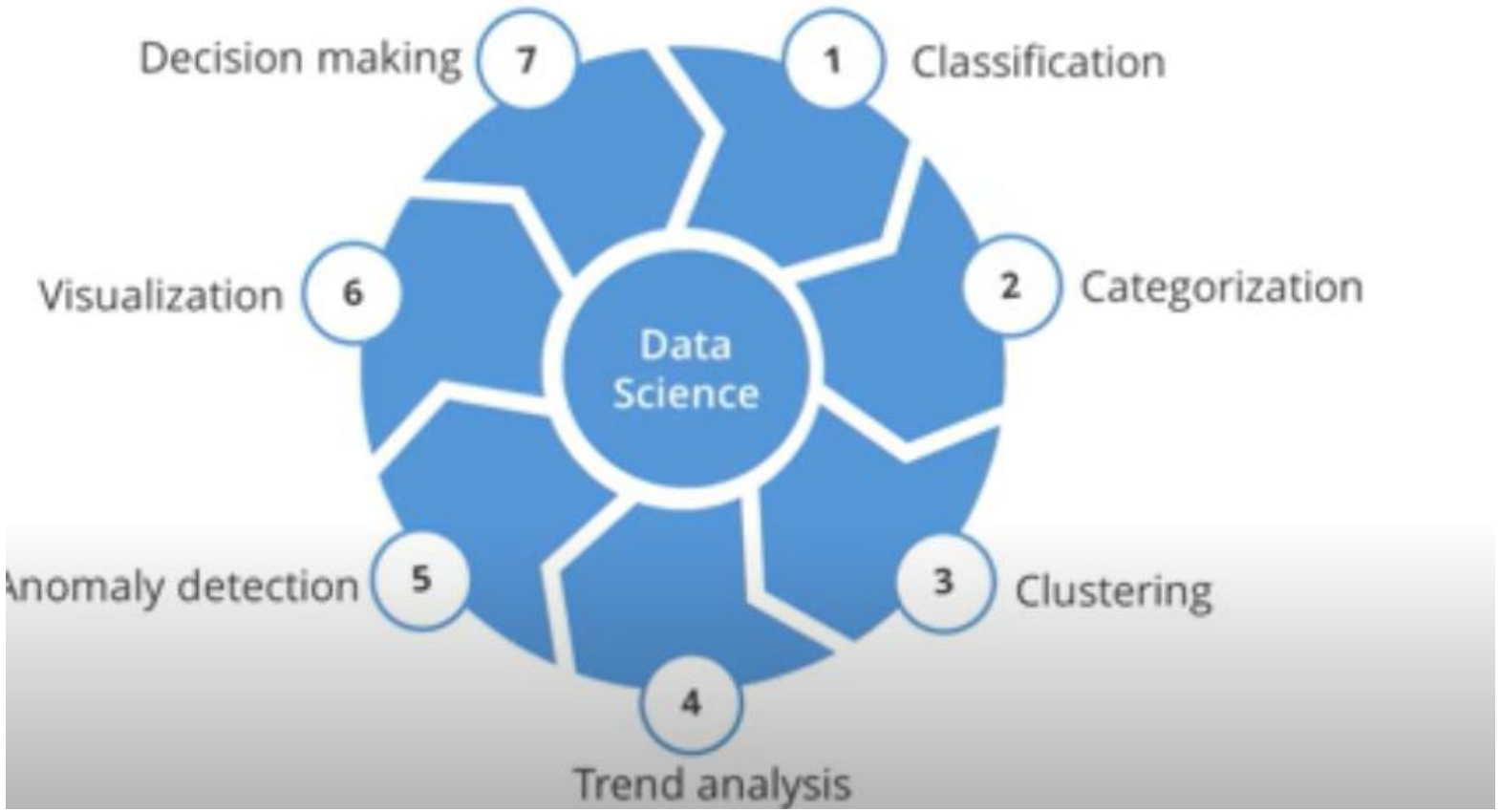

ARTIFICIAL INTELLIGENCE
A program that can sense, reason, act, and adapt
MACHINE LEARNING
Algorithms whose performance improve as they are exposed to more data over time
DEEP LEARNING
Subset of machine learning in which multilayered neural networks learn from vast amounts of data

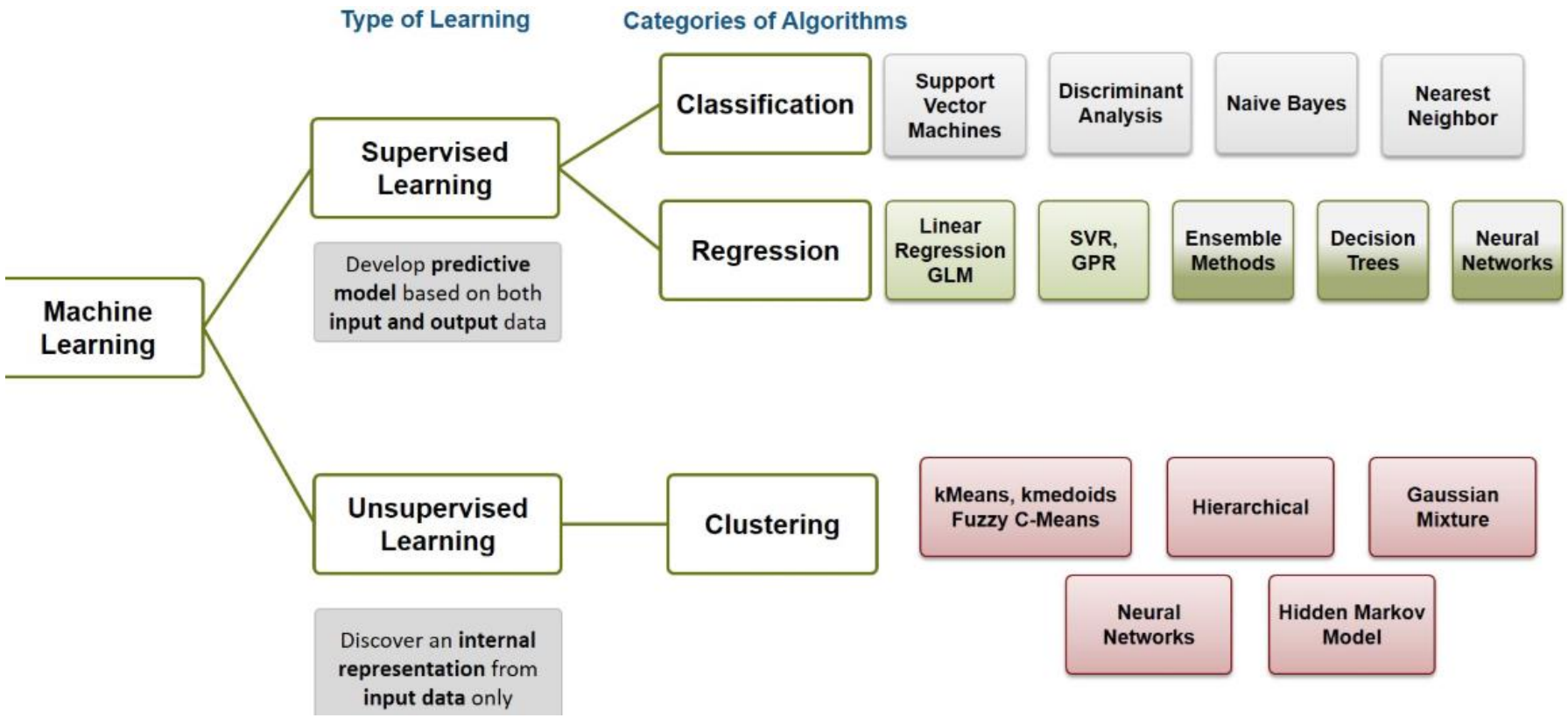
Artificial intelligence is defined as study of rational agents. A rational agent could be anything which makes decisions, like a person, firm, machine, or software. It carries out an action with the best outcome after considering past and current percepts



Machine Learning is the learning in which machine can learn by its own without being explicitly programmed. It is an application of AI that provide system the ability to automatically learn and improve from experience.




Decision making
7
1
Classification
Visualization
6
Data Science
2
Categorization
Anomaly detection
5
3
Clustering
4
Trend analysis


Type of Learning
Categories of Algorithms
Machine Learning
Supervised Learning
Develop predictive model based on both input and output data
Classification
Support Vector Machines
Discriminant Analysis
Naive Bayes
Nearest Neighbor
Regression
Linear Regression GLM
SVR, GPR
Ensemble Methods
Decision Trees
Neural Networks
Unsupervised Learning
Discover an internal representation from input data only
Clustering
kMeans, kmedoids Fuzzy C-Means
Hierarchical
Gaussian Mixture
Neural Networks
Hidden Markov Model

Applications of ML

Retail	Marketing	Healthcare	Telco	Finance
• Demand forecasting • Supply chain optimization • Pricing optimization • Market segmentation and targeting • Recommendations	• Recommendation engines & targeting • Customer 360 • Click-stream analysis • Social media analysis • Ad optimization	• Predicting Patient Disease Risk • Diagnostics and Alerts • Fraud	• Customer churn • System log analysis • Anomaly detection • Preventative maintenance • Smart meter analysis	• Risk Analytics • Customer 360 • Fraud • Credit scoring

Anaconda

Anaconda is a distribution of packages built for data science.

- It comes with conda, a package and environment manager.
- Use conda to create environments for isolating your projects that use different versions of Python and/or different packages.
- Use it to install, uninstall, and update packages in your environments.
- Anaconda is available for Windows, Mac OS X, and Linux. You can find the installers and installation instructions at **https://www.anaconda.com/download/.**

Anaconda

On Windows

A bunch of applications are installed along with Anaconda:

Anaconda Navigator, a GUI for managing your environments and packages

Anaconda Prompt, a terminal where you can use the command line interface to manage your environments and packages

Spyder, an IDE geared toward scientific development

To avoid errors later, it's best to update all the packages in the default environment. Open the Anaconda Prompt application. In the prompt, run the following commands:

.

Jupyter notebooks

The notebook is a web application that allows you to combine explanatory text ,math equations, code, and visualizations all in one easily sharable document.

- **Installing Jupyter Notebook**

  conda install jupyter notebook.

- **Launching the notebook server**

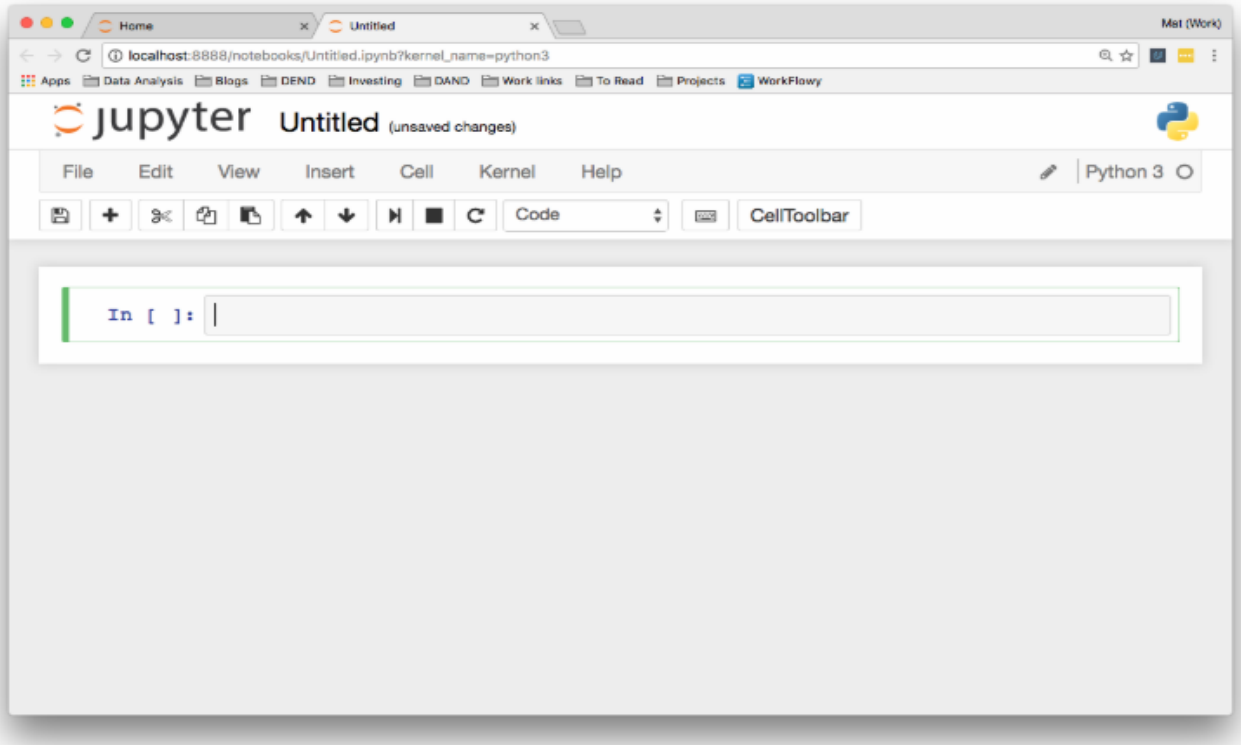
  jupyter notebook in your terminal or console.

When you run the command (try it yourself!), the server home should open in your browser. By default, the notebook server runs at **http://localhost:8888**

Jupyter notebooks

Notebook interface

When you create a new notebook, you should see something like this:



Numpy

NumPy stands for *Numerical Python* and it's a fundamental package for scientific computing in Python. NumPy provides Python with an extensive math library capable of performing numerical computations effectively and efficiently.

- NumPy is included with **Anaconda**. If you don't already have Anaconda installed on your computer.

- NumPy arrays being memory-efficient and from optimized algorithms used by NumPy for doing arithmetic, statistical, and linear algebra operations.

Pandas

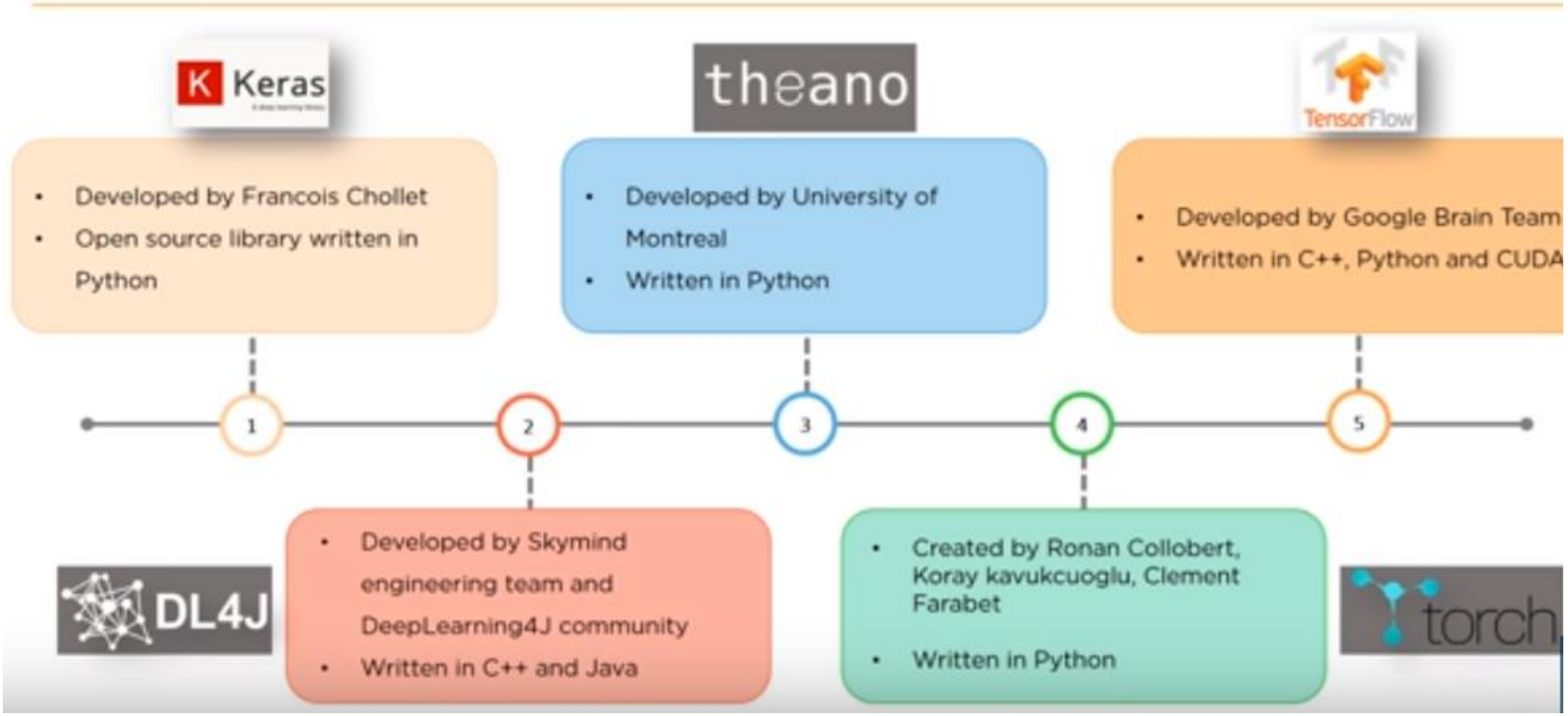
- **Pandas** is a package for data manipulation and analysis in Python.
- The name Pandas is derived from the econometrics term *Panel Data*.
- Pandas incorporates two additional data structures into Python, namely **Pandas Series** and **Pandas DataFrame**.
- These data structures allow us to work with *labeled* and *relational* data in an easy and intuitive manner

Matplotlib,Seaborn

Matplotlib is a plotting library for the Python programming language and its numerical mathematics extension NumPy.

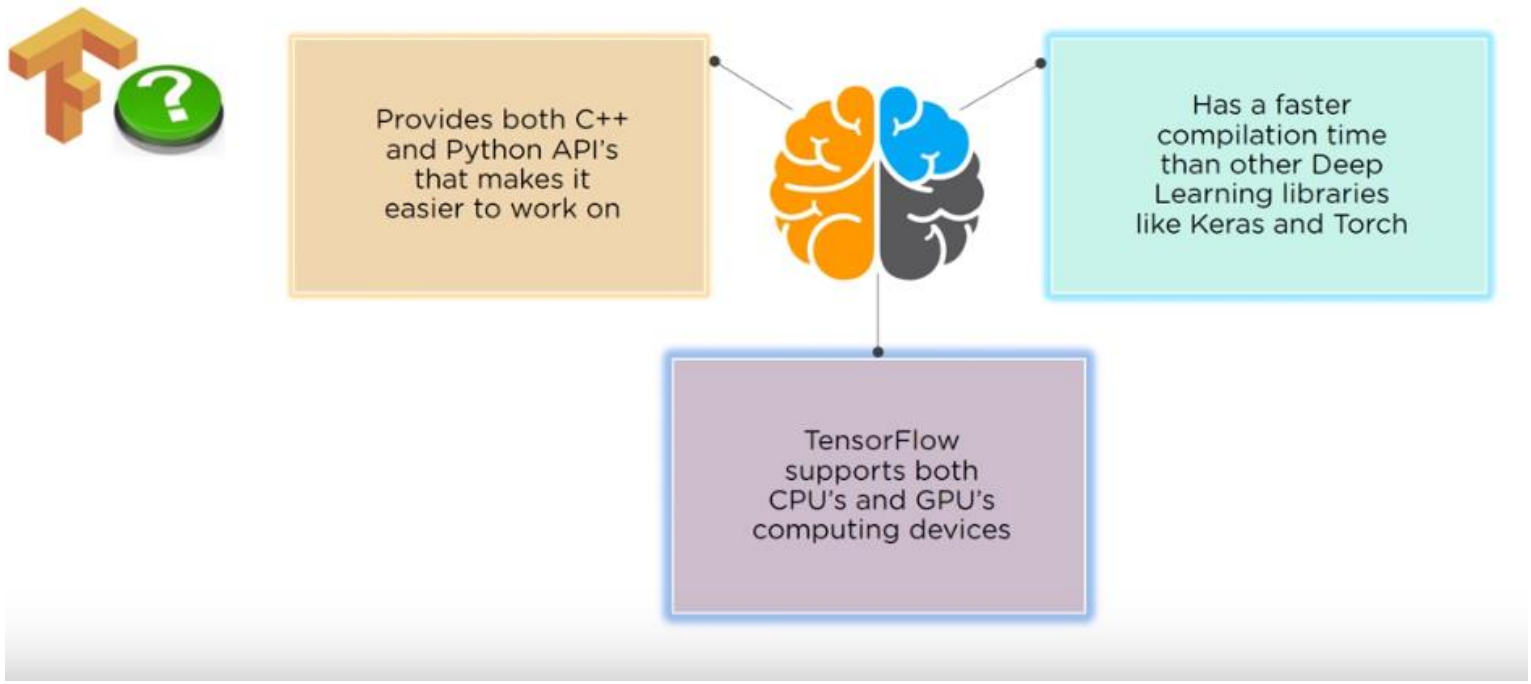
Seaborn is a Python data visualization library based on matplotlib. It provides a high-level interface for drawing attractive and informative statistical graphics.

Deep Learning Libraries



Tensorflow

TensorFlow is an open source software library created by Google for numerical computation using data flow graphs. Nodes in the graph represent mathematical operations, while the graph edges represent the multidimensional data arrays (tensors) that flow between them.



Linear Regression

Predict house value

```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
import seaborn as sb
% matplotlib inline
```

```
UsageError: Line magic function `%` not found.
```

```
#import house data
house_data= pd.read_csv('C:\\Users\\tejas\\Downloads\\kc_house_data.csv')
```

```
house_data.shape
```

```
(21613, 21)
```

Linear Regression

Predict house value

```
house_data.head()
```

	id	date	price	bedrooms	bathrooms	sqft_living	sqft_lot	floors	waterfront	view	...	grade	sqft_abov
0	7129300520	20141013T000000	221900.0	3	1.00	1180	5650	1.0	0	0	...	7	118
1	6414100192	20141209T000000	538000.0	3	2.25	2570	7242	2.0	0	0	...	7	217
2	5631500400	20150225T000000	180000.0	2	1.00	770	10000	1.0	0	0	...	6	77
3	2487200875	20141209T000000	604000.0	4	3.00	1960	5000	1.0	0	0	...	7	105
4	1954400510	20150218T000000	510000.0	3	2.00	1680	8080	1.0	0	0	...	8	168

5 rows × 21 columns

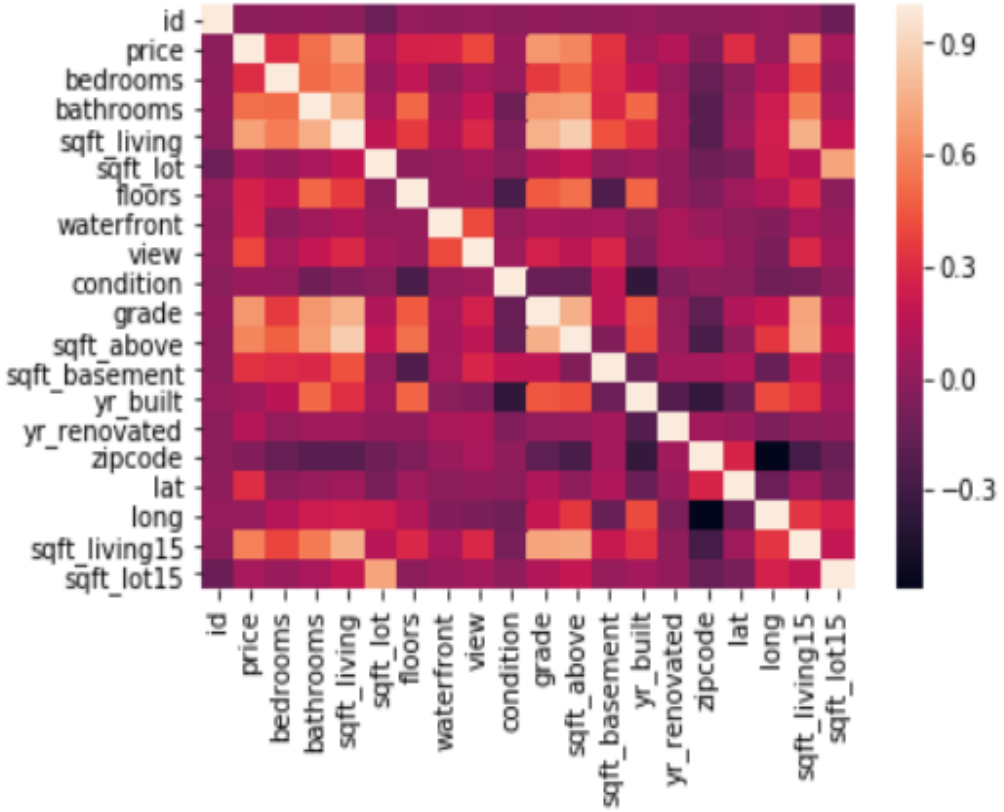
```
X= house_data.iloc[:,3:]
Y= house_data['price']
```

Linear Regression

Predict house value

```
sb.heatmap(house_data.corr())
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x20146b60dd8>
```



Linear Regression

Predict house value

```
#generating train and test data sets
from sklearn.model_selection import train_test_split
X_train,X_test,Y_train,Y_test = train_test_split(X,Y,test_size =0.2)
```

```
#define model
from sklearn.linear_model import LinearRegression
regressor = LinearRegression()
regressor.fit(X_train,Y_train)
```

```
LinearRegression(copy_X=True, fit_intercept=True, n_jobs=None,
         normalize=False)
```

```
#predicting test set results
Y_pred =regressor.predict(X_test)
Y_pred
```

```
array([888950.05175903, 499470.41047668, 731170.19770784, ...,
       980860.30003182, 542045.15156554, 651925.62361042])
```

Linear Regression

Predict house value

```
#calculating coefficients
print(regressor.coef_)
```

```
[-3.52363516e+04  4.11251490e+04  1.07949110e+02  1.59799606e-01
  1.11371110e+04  5.34400229e+05  5.31001890e+04  2.86912304e+04
  9.46108875e+04  7.00431386e+01  3.79059708e+01 -2.60329644e+03
  2.07562585e+01 -5.50145200e+02  6.03193839e+05 -2.07016515e+05
  2.15671973e+01 -4.15936350e-01]
```

```
#calculating the intercept
print(regressor.intercept_)
```

```
4412861.939381944
```

Linear Regression

Predict house value

```
#calculating r squared value
from sklearn.metrics import r2_score
r2_score(Y_test,Y_pred)
```

```
0.7076852441584968
```