Machine Learning

Practical File

BACHELOR OF TECHNOLOGY

Information Technology

SUBMITTED BY : ANJALI SINGH

CRN : 2121009

URN : 2104465

GURU NANAK DEV ENGINEERING COLLEGE

LUDHIANA-141006, INDIA


CONTENT

7 MORE TERMS TO KNOW

7.1 Simple linear regression

7.2 Multilple linear regression

7.3 Polynomial regression

7.4 SVR

7.5 Decision tree

7.6 Random Forest

7.7 KNN

7.8 SVM

7.9 K-means

7.10 K-modes

11/07/2022

1. BASICS PYTHON PROGRAM

1.1 To print name, age and roll number

CODE:

```
a=18.10
print("NAME-ANJALI SINGH")
print("CRN-2121009")
print(a)
```

OUTPUT:

1.2 To use if else statement in python

CODE:

```
a=24
if a==2:
print("TRUE")
else:
print("FALSE")
```

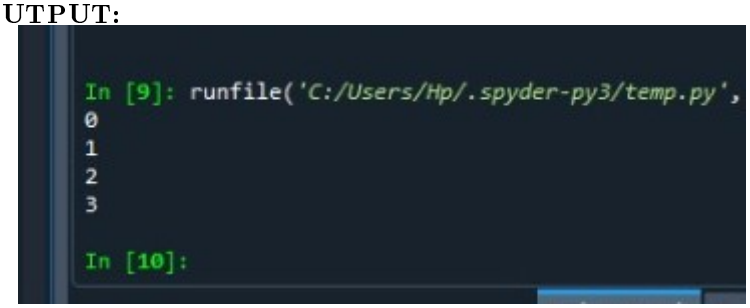
OUTPUT:

1.3 To use for loop in python

CODE:

```
n = 4
for i in range(0, n):
print(i)
```

OUTPUT:



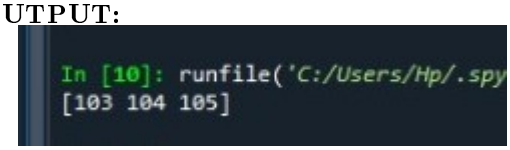
2. NUMPY LIBRARY

2.1 To use numpy library to add two numbers

CODE:

```
import numpy as np
a = np.array([98,99,100])
b = np.array([5])
c = np.add(a,b)
print(c)
```

OUTPUT:



2.2 To use Numpy Library to find the min marks from an array

CODE:

```
import numpy as np
student_marks_arr = np.array([78, 92, 36, 64, 89])
student_marks_arr
student_marks_arr.shape
a=min(student_marks_arr)
print(a)
```

OUTPUT:

2.3 To use Numpy Libray to find the maximum car attribute from an array CODE:

```
import numpy as np
car_attributes = [[18, 15, 18, 16, 17],[130, 165, 150, 150, 140],[307, 350, 318, 304, 302]]
car_attributes_arr = np.array(car_attributes)
a=np.max(car_attributes_arr)
print(a)
```

OUTPUT:

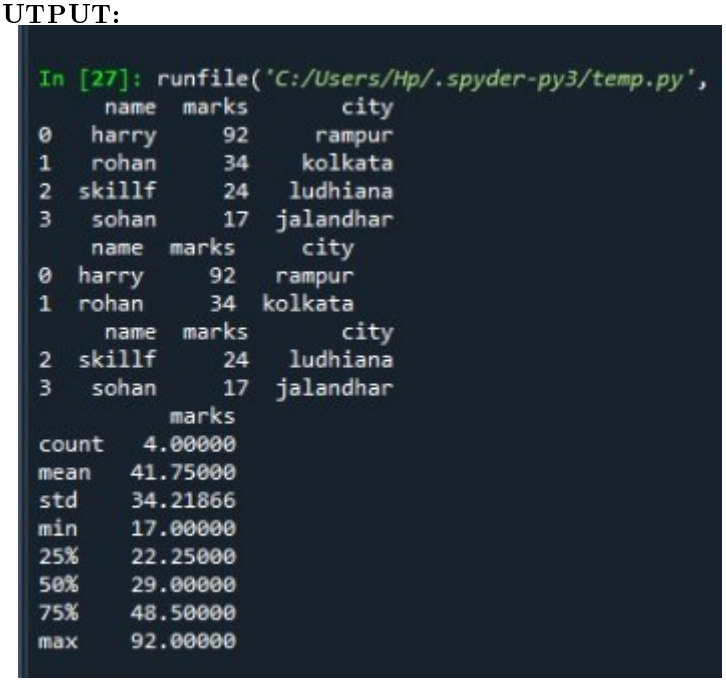
3. PANDA LIBRARY

3.1 Using Panda for data framing

CODE:

```
import pandas as pd
dic1= {
"name": ['harry','rohan','skillf','sohan'],
"marks":[92,34,24,17], "city":['rampur','kolkata','ludhiana','jalandhar']
}
df=pd.DataFrame(dic1)
print(df)
df.to_csv('a.csv')
print(df.head(2))
print(df.tail(2))
print(df.describe())
df['marks']
df['marks'][0]
```

OUTPUT:



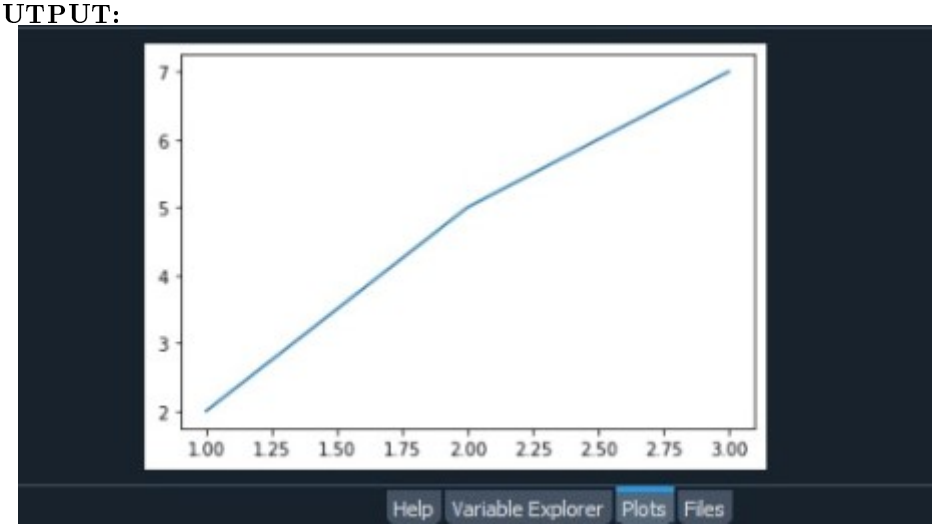
4. MATPLOTLIP

To use matplotlib library

CODE:

```
import matplotlib.pyplot as plt
plt.plot([1,2,3],[2,5,7])
plt.show()
```

OUTPUT:



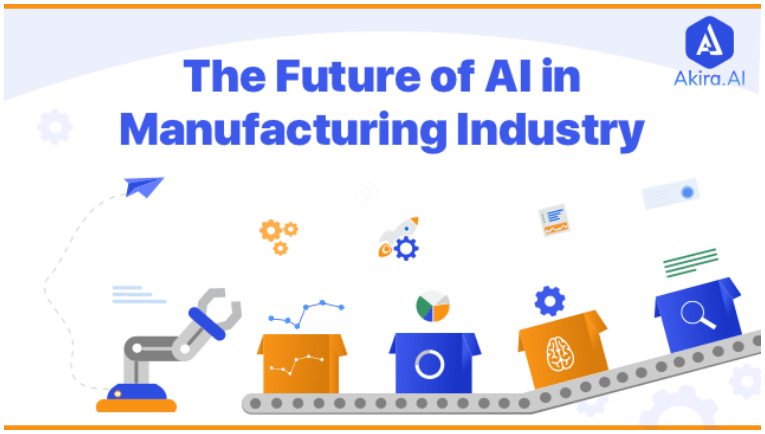
12/07/2022

5.REPORT ON

ARTIFICIAL INTELLIGENCE IN MANUFACTURING

AI is a pervasive technology, the application and impact of which span all sectors of the economy, from manufacturing and media to education to healthcare. The four types of intelligences are as follows (Huang & Rust, 2018):

1. Analytical (learns and adapts systematically based on data) .
2. Intuitive (learns and adapts intuitively based on understanding).
3. Empathetic (learns and adapts empathetically based on experience.
4. Machanics (minimal degree of learning or adaptation).



5.1. Introduction:

Intelligent manufacturing is a constantly evolving concept, which can be summarized into three basic prototypes: digital manufacturing , digital-networked manufacturing (2nd generation), and new-generation intelligent manufacturing . By analysing the evolution of intelligent manufacturing, it is evident, that the process of developing from traditional manufacturing to intelligent manufacturing is also a process of developing from the original human-physical systems (HPS) to human-cyber-physical systems (HCPS). An HCPS reveals the basic principles of intelligent manufacturing development and is the theoretical basis for supporting the development of new-generation intelligent manufacturing.

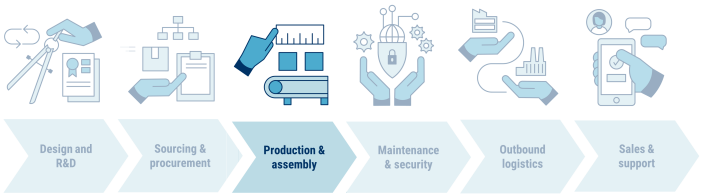
5.2. How is AI used in the manufacturing industry?

In manufacturing, AI can oer the most value in planning and production oor operations. According to a BCG report, the most important AI use cases in the manufacturing industry are:

1. Intelligent
2. self-optimizing machines that automate production processes
3. Forecasting eciency losses for better planning
4. Detecting quality defects to facilitate predictive maintenance

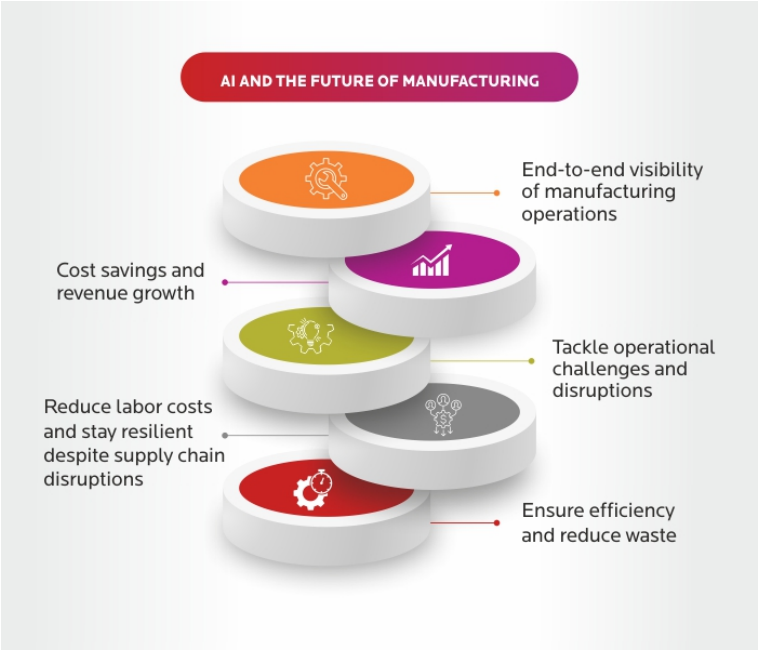
5.3. How will AI change manufacturing?

The emergence of 4IR technologies will usher in the era of smart manufacturing with digital factories. According to the IFR (International Federation of Robotics), there were already 2.7 million industrial robots operating in factories worldwide in 2020. This was a 12% increase as compared to 2019, and with digitization initiatives growing exponentially, the trend is projected to grow further. Manufacturers will continue to invest in technologies like AI and machine learning to further bring down production costs and improve time-to-market. In the aftermath of a global pandemic, manufacturers will strive to make their businesses more resilient by adopting technologies that automate tasks, forecast disruptions, and facilitate end-to-end control of all operations.



5.4. AI In Manufacturing Industry |Use Cases

1. Articial intelligence in logistics
2. AI robots
3. Articial intelligence in supply chain management
4. AI autonomous vehicles
5. AI for factory automation
6. AI for IT operations
7. AI in design and manufacturing
8. Articial intelligence and IoT
9. AI in warehouse management
10. AI process automation
11. AI for predictive maintenance
12. AI-based product development
13. AI-based connected factory
14. AI-based visual inspections and quality control
15. AI for purchasing price variance
16. AI order management
17. AI for cybersecurity



There are many areas where artical intelligence and machine learning are transforming manufacturing today like:

*Machine maintenance:

Modern maintenance management systems aim to ensure the optimal, complete, and safe operation of all critical facilities and production equipment. Machine learning in manufacturing can help by implementing ecient predictive maintenance for critical production equipment.

*Cybersecurity:

With the drastic rise of cyber threats, companies are looking for new ways to defend themselves from these attacks. As these challenges increase over time, many organizations are turning to AI to help solve their most signicant cybersecurity challenges. With large amounts of data and operating technology environments producing mountains of logs and dataand their respective networks, appliances, and applicationsAI is the perfect solution, since it's able to process and analyze these enormous volumes of information.

*Robotics:

The future of AI in manufacturing is largely dependent on the development of new robotics. Industrial robots, also called manufacturing robots, are indispensable tools that can be used to increase eciency and productivity. They can help manage human workers by performing repetitive tasks, such as packing, shipping, and inspection. Industrial robots are also used in all types of industries, including appliance assembly, food processing, injection molding, metalworking, welding, and painting.

5.5. What are the benets of AI in Manufacturing?

1. Safety:

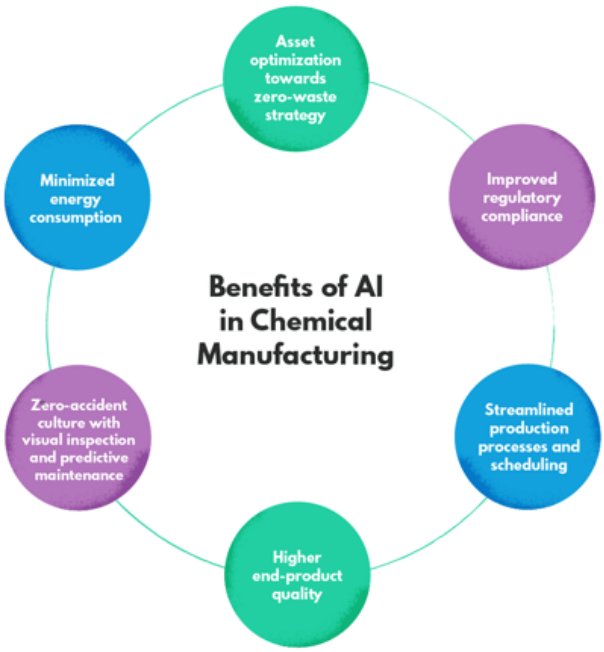
As one of the most dangerous industries, there are more than 3,000 major injuries and nine fatalities in manufacturing each year in the United States, which shows how important it is for companies to use AI-powered robots to reduce accidents in their work areas.

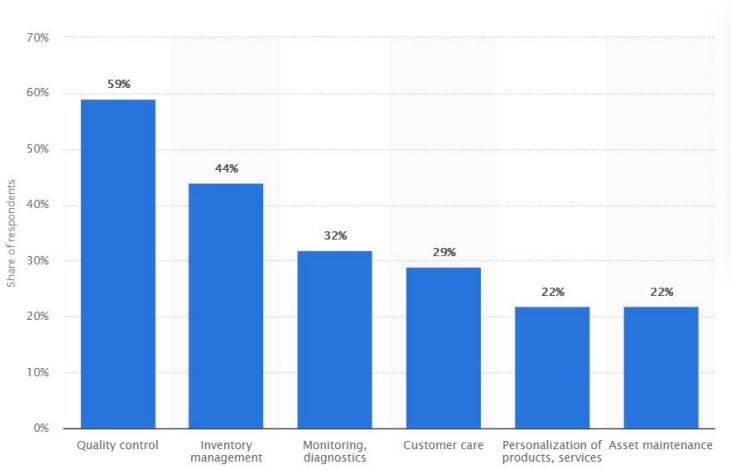
2. Eciency:

An important aspect of AI implementation in the manufacturing industry is the ability to produce at any time, regardless of human limitations. Robots can work around the clock, which means that businesses can increase production by a huge amount. They'll also be able to scale for mass consumers.

3. Cost reduction:

An additional benet of AI in manufacturing is its ability to reduce the operation costs of manufacturers. Leveraging AI technologies can enhance organizations' analytics capabilities and thus allow them to reduce the use of resources, better predict the future, and boost their productivity. The use of AI can also lead to more accurate sales and production forecasts, reduced waste in inventory, and improved nancial results.





5.6. Conclusion:

The future of AI in manufacturing is looking bright, with machine learning improving eciency in manufacturing work environmentsand leading to better industry performance overall. For years robotics, advanced analytics, and automation has been a major part of the manufacturing industry. The increasing scale of adoption of AI in manufacturing seems more like an evolution, rather than an industry disruption. Technology is already here and more massive implementation is a matter of time.

13/07/2022

6.REPORT ON AI TOOLS

6.1.SCIKIT LEARN

6.1.1.Overview

The scikit-learn project started as scikits.learn, a Google Summer of Code project by French data scientist David Cournapeau. Its name stems from the notion that it is a "SciKit" (SciPy Toolkit), a separately-developed and distributed third-party extension to SciPy.The original codebase was later rewritten by other developers. In 2010 Fabian Pedregosa, Gael Varoquaux, Alexandre Gramfort and Vincent Michel, all from the French Institute for Research in Computer Science and Automation in Rocquencourt, France, took leadership of the project and made the first public release on February the 1st 2010.

6.1.2.FEATURES OF SCIKIT LEARNING:-

1. DATASETS: Scikit-learn comes with several inbuilt datasets such as the iris dataset, house prices dataset, diabetes dataset, etc. The main functions of these datasets are that they are easy to understand and you can directly implement ML models on them.

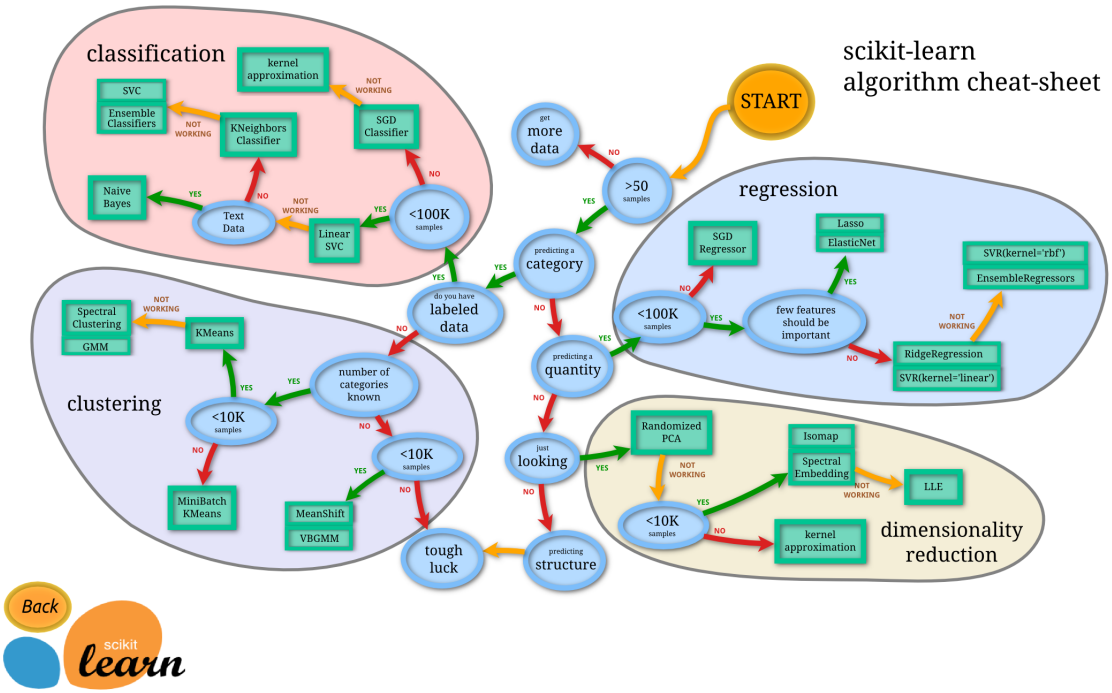
2. DATA SPLITTING : Sklearn provided the functionality to split the dataset for training and testing. Splitting the dataset is essential for an unbiased evaluation of prediction performance. We can define what proportion of our data to be included in train and test datasets.

3. LINEAR REGRESSION : This supervised ML model is used when the output variable is continuous and it follows linear relation with dependent variables. It can be used to forecast sales in the coming months by analyzing the sales data for previous months.

4. DECISION TREE : A Decision Tree is a powerful tool that can be used for both classification and regression problems. It uses a tree-like model to make decisions and predict the output. It consists of roots and nodes. Roots represent the decision to split and nodes represent an output variable value.

5. BOOSTING : Boosting is a technique in which multiple models are trained in such a way that the input of a model is dependent on the output of the previous model. In Boosting, the data which is predicted incorrectly is given more preference. Ex- Ada Boost, Gradient Boost.

6. RANDOM FOREST : Random Forest is a bagging technique in which hundreds/thousands of decision trees are used to build the model. Random Forest can be used for both classification and regression problems. It can be used to classify loan applicants, identify fraudulent activity and predict diseases.



6.2.APACHE MXNET

6.2.1.OVERVIEW

Apache MXNet is an open-source deep learning software framework, used to train and deploy deep neural networks. It is scalable, allowing for fast model training and supports a flexible programming model and multiple programming languages (including C++, Python, Java, Julia, MATLAB, JavaScript, Go, R, Scala, Perl, and Wolfram Language). The MXNet library is portable and can scale to multiple GPUs as well as multiple machines. It was co-developed by Carlos Guestrin at University of Washington.

6.2.2.FEATURES APACHE MXNET is a scalable deep learning framework that supports deep learning models, such as; convolutional neural networks (CNNs) and long short-term memory networks (LSTMs).

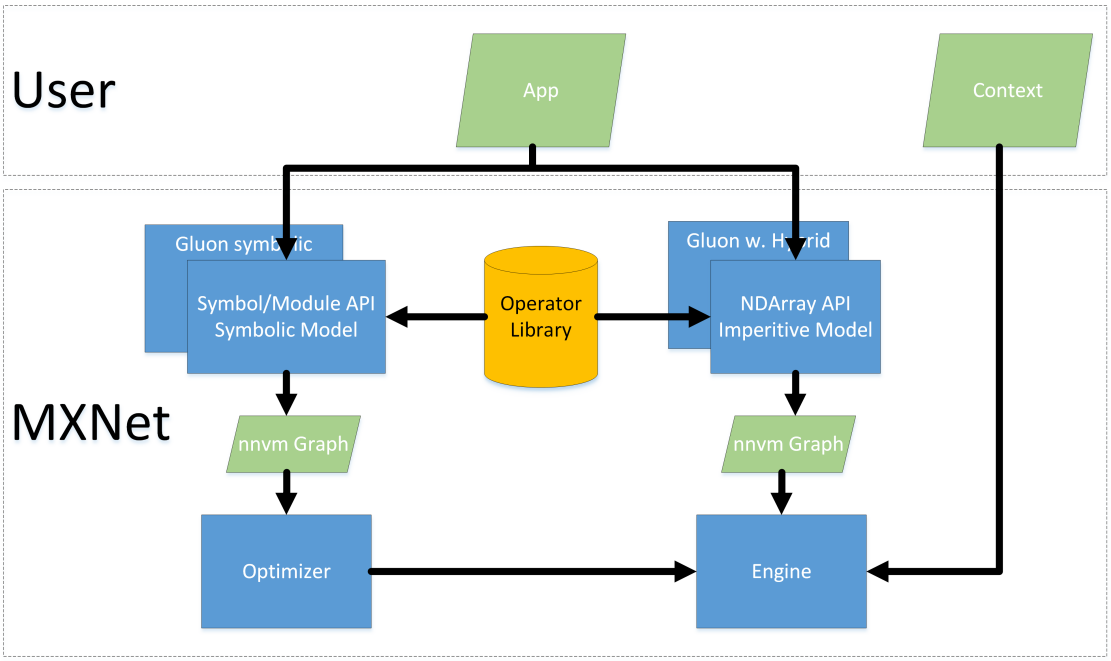
1. SCALABLE : MXNet can be distributed on dynamic cloud infrastructure using a distributed parameter server (based on research at Carnegie Mellon University, Baidu, and Google[4]). with multiple GPUs or CPUs the framework approaches linear scale.

2. FLEXIBLE : MXNet supports both imperative and symbolic programming. The framework allows developers to track, debug, save checkpoints, modify hyperparameters, and perform early stopping.

3. MULTIPLE LANGUAGES : MXNet supports Python, R, Scala, Clojure, Julia, Perl, MATLAB and JavaScript for front-end development, and C++ for back-end optimization.

4. PORTABLE : Supports an efficient deployment of a trained model to low-end devices for inference, such as mobile devices (using Amalgamation[5]), Internet of things devices (using AWS Greengrass), serverless computing (using AWS Lambda) or containers. These low-end environments can have only weaker CPU or limited memory (RAM), and should be able to use the models that were trained on a higher-level environment (GPU based cluster, for example).

5. CLOUD SUPPORT : MXNet is supported by public cloud providers including Amazon Web Services (AWS)[6] and Microsoft Azure.[7] Amazon has chosen MXNet as its deep learning framework of choice at AWS.


User
App
Context
Gluon symbolic
Symbol/Module API
Symbolic Model
Operator
Library
Gluon w. Hybrid
NDArray API
Imperitive Model
MXNet
nnvm Graph
nnvm Graph
Optimizer
Engine

6.3.TENSORFLOW

6.3.1OVERVIEW

TensorFlow is a free and open-source software library for machine learning and artificial intelligence. It can be used across a range of tasks but has a particular focus on training and inference of deep neural networks
TensorFlow was developed by the Google Brain team for internal Google use in research and production.The initial version was released under the Apache License 2.0 in 2015. Google released the updated version of TensorFlow, named TensorFlow 2.0, in September 2019.

6.3.2FEATURES OF TENSORFLOW:-

1. AUTODIFFERENTIATION : AutoDifferentiation is the process of automatically calculating the gradient vector of a model with respect to each of its parameters. With this feature, TensorFlow can automatically compute the gradients for the parameters in a model, which is useful to algorithms such as backpropagation which require gradients to optimize performance.[32] To do so, the framework must keep track of the order of operations done to the input Tensors in a model, and then compute the gradients with respect to the appropriate parameters.

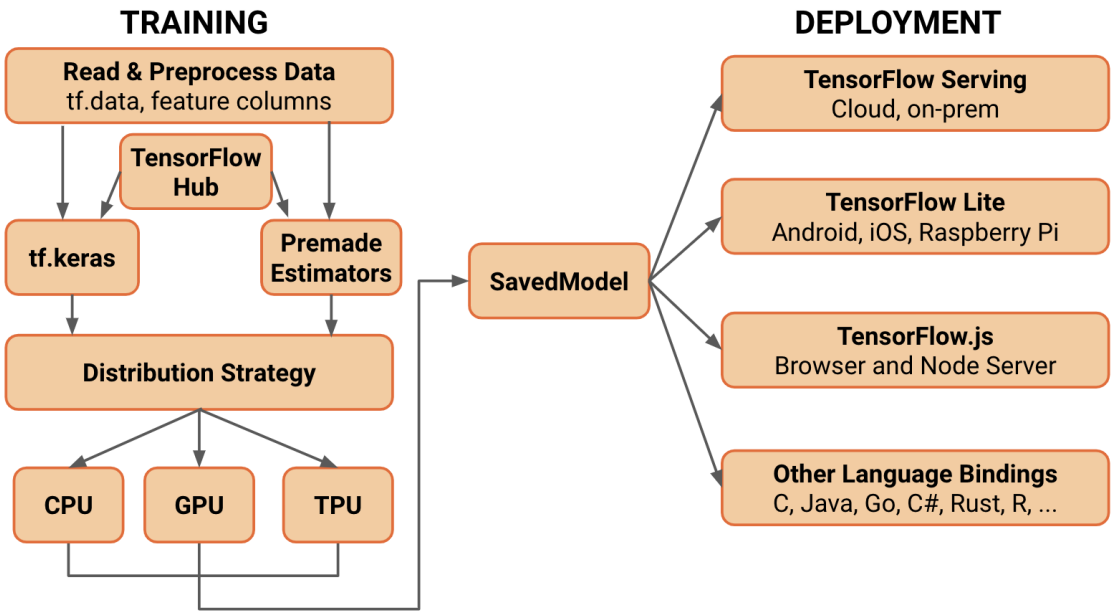
2. EAGER EXECUTION : TensorFlow includes an "eager execution" mode, which means that operations are evaluated immediately as opposed to being added to a computational graph which is executed later.[33] Code executed eagerly can be examined step-by step-through a debugger, since data is augmented at each line of code rather than later in a computational graph.[33] This execution paradigm is considered to be easier to debug because of its step by step transparency.

3. DISTRIBUTE : In both eager and graph executions, TensorFlow provides an API for distributing computation across multiple devices with various distribution strategies.[34] This distributed computing can often speed up the execution of training and evaluating of TensorFlow models and is a common practice in the field of AI.

4. LOSSES : To train and assess models, TensorFlow provides a set of loss functions (also known as cost functions).[36] Some popular examples include mean squared error (MSE) and binary cross entropy (BCE).[36]

These loss functions compute the "error" or "difference" between a model's output and the expected output (more broadly, the difference between two tensors). For different datasets and models, different losses are used to prioritize certain aspects of performance.

5. METRICS : In order to assess the performance of machine learning models, TensorFlow gives API access to commonly used metrics. Examples include various accuracy metrics (binary, categorical, sparse categorical) along with other metrics such as Precision, Recall, and Intersection-over-Union .

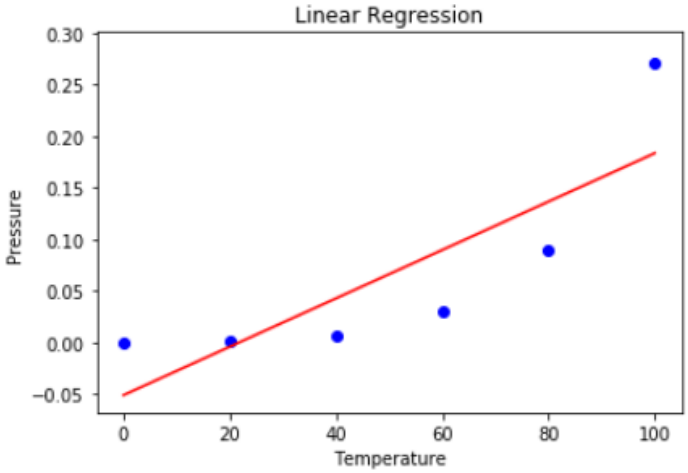


14/07/2022

7.More Terms To Know

7.1.Simple Linear Regression

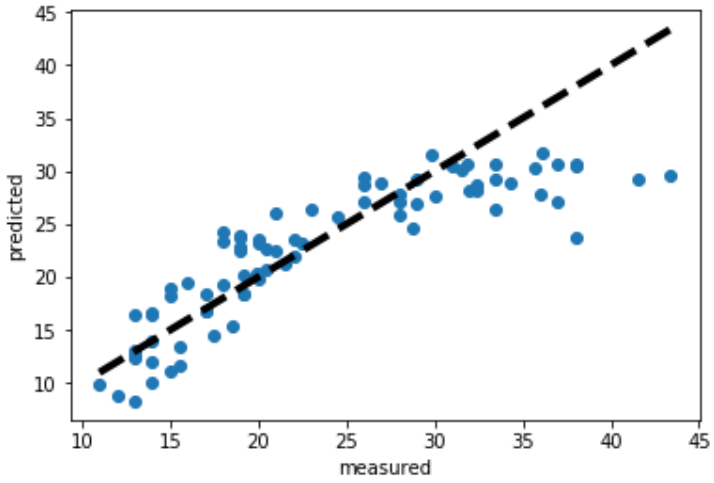
Simple linear regression is a statistical method that we can use to find a relationship between two variables and make predictions. The two variables used are typically denoted as y and x. The independent variable, or the variable used to predict the dependent variable is denoted as x. The dependent variable, or the outcome/output, is denoted as y.



7.2.Multiple Linear Regression

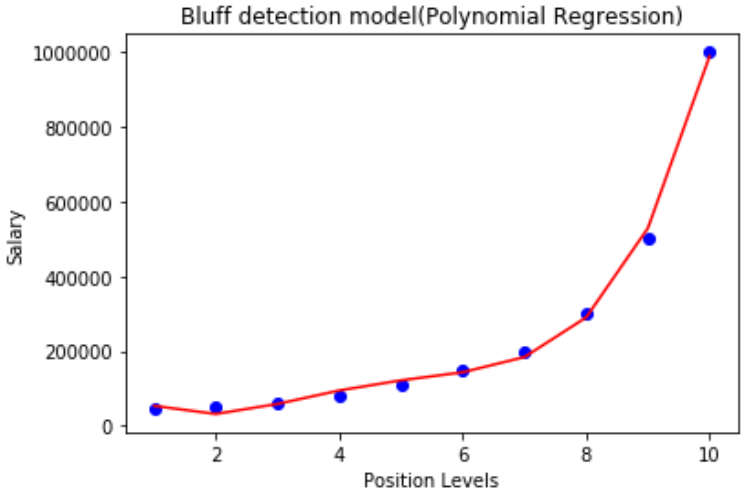
Multiple linear regression attempts to model the relationship between two or more features and a response by fitting a linear equation to the observed data.

Clearly, it is nothing but an extension of simple linear regression.



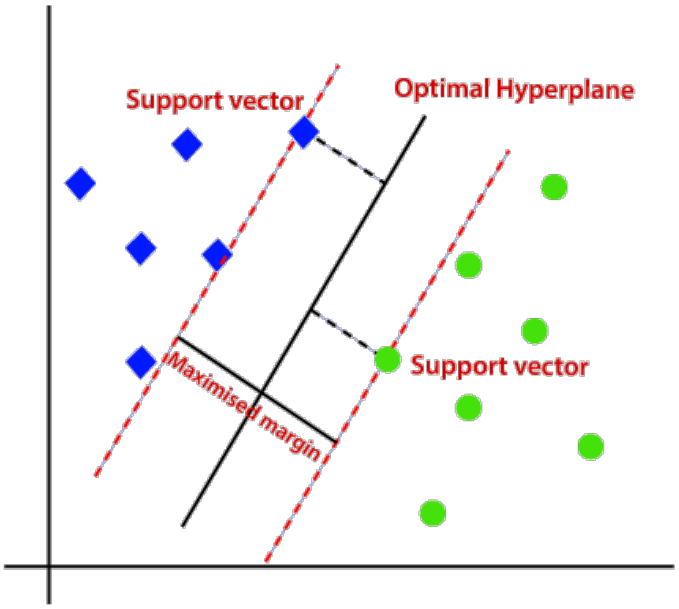
7.3. Polynomial Linear Regression

Polynomial Regression is a form of linear regression in which the relationship between the independent variable x and dependent variable y is modeled as an nth degree polynomial.



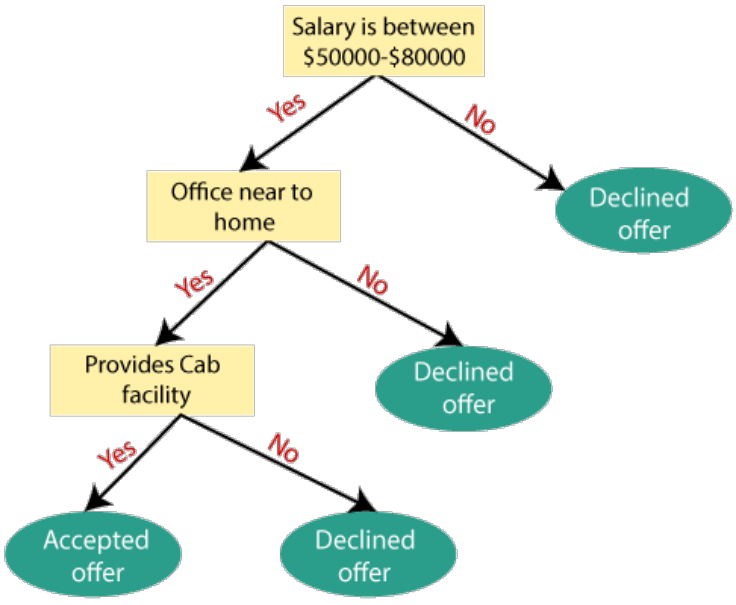
7.4. Support Vector Regression(SVM)

Support Vector Regression is a supervised learning algorithm that is used to predict discrete values.



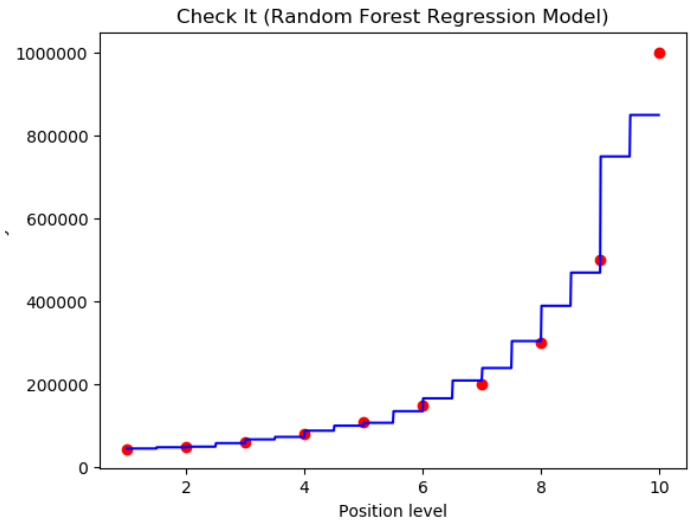
7.5. Decision Tree

Decision Tree is one of the most commonly used, practical approaches for supervised learning. It can be used to solve both Regression and Classification tasks with the latter being put more into practical application.



7.6. Random Forest Regression

Random Forest Regression is a supervised learning algorithm that uses ensemble learning method for regression.


Check It (Random Forest Regression Model)
1000000
800000
600000
400000
200000
0
2
4
6
8
10
Position level

15/07/2022

7.7.KNN

k-nearest neighbor algorithm:

This algorithm is used to solve the classification model problems. K-nearest neighbor or K-NN algorithm basically creates an imaginary boundary to classify the data. When new data points come in, the algorithm will try to predict that to the nearest of the boundary line.

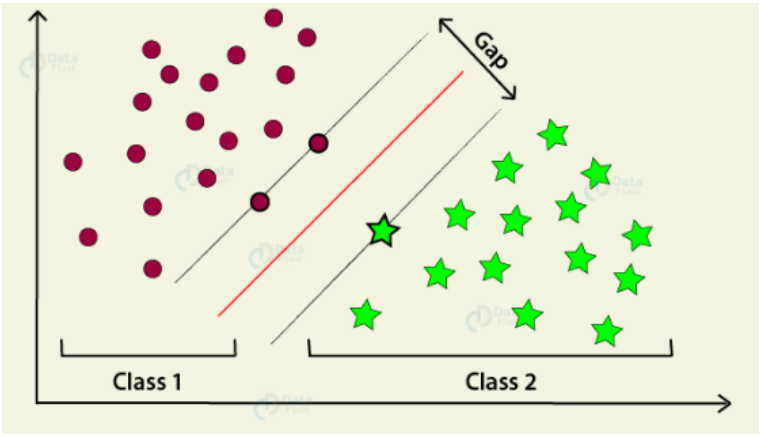
CODE:

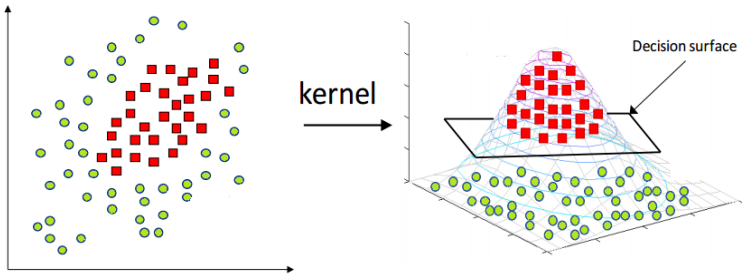
```
# Import necessary modules
from sklearn.neighbors import KNeighborsClassifier
from sklearn.model_selection import train_test_split
from sklearn.datasets import load_iris
import numpy as np
import matplotlib.pyplot as plt
irisData = load_iris()
# Create feature and target arrays
X = irisData.data
y = irisData.target
# Split into training and test set
X_train, X_test, y_train, y_test = train_test_split( X, y, test_size = 0.2,
random_state=42)
neighbors = np.arange(1, 9)
train_accuracy = np.empty(len(neighbors))
test_accuracy = np.empty(len(neighbors))
# Loop over K values
for i, k in enumerate(neighbors):
knn = KNeighborsClassifier(n_neighbors=k)
knn.fit(X_train, y_train)
# Compute training and test data accuracy
train_accuracy[i] = knn.score(X_train, y_train)
test_accuracy[i] = knn.score(X_test, y_test)
# Generate plot
plt.plot(neighbors, test_accuracy, label = 'Testing dataset Accuracy')
plt.plot(neighbors, train_accuracy, label = 'Training dataset Accuracy')
plt.legend()
plt.xlabel('n_neighbors')
plt.ylabel('Accuracy')
plt.show()
```

OUTPUT:

7.8.Support Vector Machine(SVM)

"Support Vector Machine" (SVM) is a supervised machine learning algorithm that can be used for both classification or regression challenges. However, it is mostly used in classification problems.

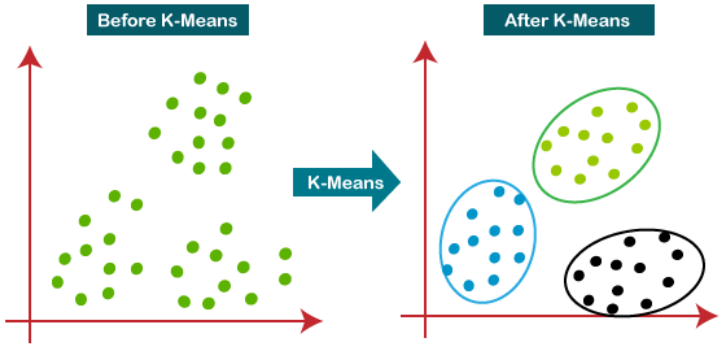




19/07/2022

7.9.K-Means Clustering

K-Means Clustering is an unsupervised learning algorithm that is used to solve the clustering problems in machine learning or data science.



7.10.K-Modes Clustering

K-Modes clustering is one of the unsupervised Machine Learning algorithms that is used to cluster categorical variables.