1. Time series analysis is a method of studying how a variable changes over time. Time series analysis can help you identify patterns, trends, cycles, seasonality, and outliers in your data, as well as forecast future values based on historical data.

Regression analysis is a method of exploring the relationship between a dependent variable and one or more independent variables.

Regression analysis can help you understand how different factors affect your outcome of interest, as well as predict the outcome for new or unseen cases.

Time series and regression are both methods of predictive analytics, but they have different assumptions, techniques, and applications.

* Time series analysis focuses on how a single variable changes over time, while regression analysis focuses on how multiple variables interact with each other.
* Time series assumes that the data is ordered and dependent on time, while regression assumes that the data is independent and random.
* Time series is more suitable for forecasting and detecting patterns in temporal data, while regression is more suitable for estimating and explaining the effect of variables on an outcome.

1. Data pre-processing is a crucial step in any machine learning project. It involves transforming raw data into a suitable format for the learning algorithms. By cleaning and formatting the data, we can ensure that the algorithm is only considering relevant information and that it is not being influenced by any irrelevant or incorrect data.

It not only improve the performance of the model but also make the model more interpretable and robust.

As handling missing values, removing outliers and scaling the data can help to prevent overfitting, which can lead to models that generalize better to new data.

1. A stationary time series is one whose properties do not depend on the time at which the series is observed. It a property of a time series in which the statistical properties of the series, such as mean, variance and autocorrelation, are constant over time. A stationary data will have a constant mean, and variance. The data will not have a predictable pattern in the long-term and no periodic fluctuations.
2. ARMA models are stationary models that use autoregressive (AR) and moving average (MA) components to predict future values from historical data.

ARIMA models are similar to ARMA models, but also include an integrated (I) component which accounts for non-stationary time series by taking a series of differences. Differencing is used to make the time series stationary, which means that the statistical properties of the series, such as mean and variance, do not change over time.

ARMA models should be used for stationary time series problems when the time series data does not have a trend or seasonal component. , while ARIMA models should be used for non-stationary time series problems.

import matplotlib

import matplotlib.pyplot as plt

import pandas as pd

from datetime import datetime

**import** matplotlib.pyplot as plt

x **=** [10, 20, 30, 40]

y **=** [20, 25, 35, 55]

plt.plot(x, y)

plt.show()

6)Recurrent Neural Networks works on the principle of saving the output of a particular layer and feeding this back to the input in order to predict the output of the layer. An RNN can handle sequential data, accepting the current input data, and previously received inputs.

RNN enable you to model time-dependent and sequential data problems, like stock exchange prediction, artificial intelligence, and text generation. RNN is tough to train due to the gradient problem. RNNs suffer from the matter of vanishing gradients. Thus training an RNN is a very difficult task.

1. The vanishing gradient problem refers to the issue of diminishing gradients during the training of deep neural networks.

It occur when the gradients become too small during backpropagation. This can happen in RNNs because they have recurrent connections that allow them to store information from previous time steps. These connections create long-term dependencies, which means that the gradient of a parameter depends on many previous inputs and outputs.

1. Long Short Term Memory in short LSTM is a special kind of RNN capable of learning long term sequences.  It is explicitly designed to avoid long term dependency problems.

Gated Recurrent Unit, GRU’s workflow is the same as the RNN but the difference is in the operation and gates associated with each GRU unit.

Both addresses vanishing and exploding gradient issue.

The difference between GRU & LSTM is:

* GRU uses less training parameter and therefore uses less memory and executes faster than LSTM whereas LSTM is more accurate on a larger dataset.
* LSTM is better at dealing with large sequences and accuracy is a concer whereas GRU is used when we want less memory consumption and want faster results.

1. A convolutional neural network (CNN) is a network architecture for deep learning that learns directly from data. CNNs are particularly useful for finding patterns in images to recognize objects, classes, and categories.

CNNs can be used to predict time series data. They can be applied to time series data by treating the data as an image with one dimension representing time and the other dimension representing the features of the time series. This approach allows the CNN to capture local patterns and dependencies within the time series data, making it suitable for tasks such as time series forecasting and anomaly detection.

1. Transfer learning is the process of pre-training a flexible model on a large dataset and using it later on other data with little to no training.

It involves training a deep learning model once on a large and diverse dataset of time series, and see that it performs competitively when used to forecast different time series, in different datasets.

This also means that time series coming from different domains (such as demographics, finance or industry) can share some common features.

Due to  transfer learning models can adapt themselves to new tasks without further training and therefore is beneficial in situations where the inference time has to be minimised.

1. The attention mechanism allows the model to "pay attention" to certain parts of the data and to give them more weight when making predictions.

* Time-series data often contains irregular patterns, such as sudden spikes or drops, seasonality, or trend shifts. By assigning varying weights to different elements of the input sequence, attention mechanisms enable models to focus on relevant information, to be flexible to adapt to changing patterns and make accurate predictions.
* Attention mechanisms ensure efficient interpretability of time-series forecasting models as users can understand which parts of the historical data are most influential in making predictions by gaining insights into the driving factors behind the forecasts, making it easier to validate and trust the model’s predictions.

12) Windowing method bases the model's predictions on consecutive sample windows, that is, It refers to the partitioning of continuous data streams into smaller, manageable subsets, or 'windows', for more efficient processing and analysis.

It allow access to data in the records right before and after the current record.

Data windowing acts as a bridge, allowing models to effectively capture temporal dependencies by segmenting historical data into meaningful windows.