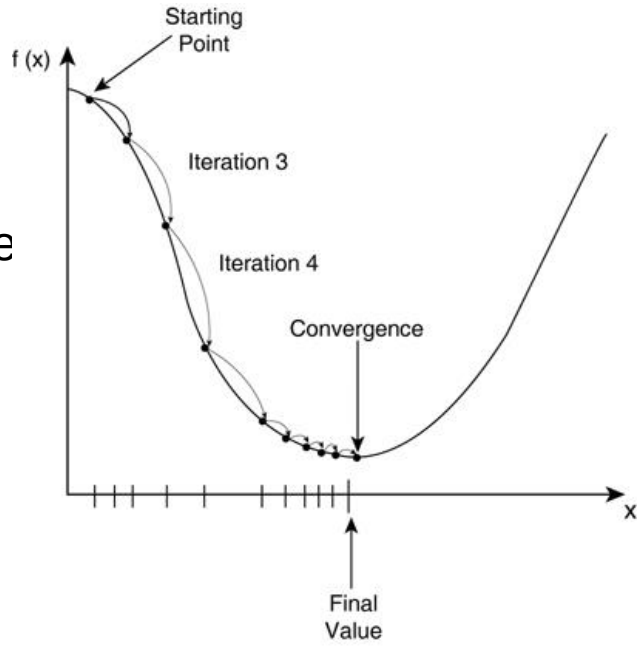
ASSIGNMENT 5

NAME: FALAK KHAN

BATCH: A3

1. How does Gradient Descent help to optimize Linear Regression model ?
>>Gradient descent enables a model to learn the gradient or direction that the model should take in order to reduce errors (differences between actual y and predicted y). ... At this point the model has optimized the weights such that they minimize the cost function. Gradient descent, therefore, enables the learning process to make corrective updates to the learned estimates that move the mode toward an optimal combination of parameters. And thus helps optimize the linear regression model effectively



2. How should a Data Professional choose a value of Learning Rate initially ?
>> A traditional default value for the learning rate is 0.1 or 0.01, and this may represent a good starting point on your problem. Given a perfectly configured learning rate, the model will learn to best approximate the function given available resources in a given number of training epochs.

Generally, a large learning rate allows the model to learn faster . A smaller learning rate may allow the model to learn a more optimal or even globally optimal set of weights but may take significantly longer to train.

3.What are the possible ways of increasing the accuracy of Linear Regression model ?

1. >> Normalise your data
2. Turn categorical data into variables via, e.g., OneHotEncoding
3. Do feature engineering:
    1. Are my features collinear?
    2. Do any of my features have cross terms/higher-order terms?
4. Regularisation of the features to reduce possible overfitting
5. Look at alternative models given the underlying features and the aim of the project

4. What is the difference between R-squared and adjusted R-squared ?

1. R-squared increases, even if the independent variable is insignificant. It never declines. Whereas Adjusted R-squared increases only when independent variable is significant and affects dependent variable.
2. Adjusted r-squared can be negative when r-squared is close to zero.
3. Adjusted r-squared value always be less than or equal to r-squared value.

Variables	R-Squared	Adjusted R- Squared
1	67.5	67.1
2	85.9	84.2
3	88.9	81.7

5. What is a Q-Q plot in Linear Regression ?

>> Quantile-Quantile (Q-Q) plot, is a graphical tool to help us assess if a set of data plausibly came from some theoretical distribution such as a Normal, exponential or Uniform distribution.

A q-q plot is a plot of the quantiles of the first data set against the quantiles of the second data set.