SADIQ TECH SOLUTIONS

DATA SCIENCE WITH AI

WHERE DATA MEETS ARTIFICIAL INTELLIGENCE

MASTER DATA, MACHINE LEARNING & AI STEP BY STEP

About Institute?

Sadiq Tech Solutions is a professional Training Organization for software developers, IT administrators, and other professionals. It's Located in Hyderabad, India. The training is offered in Four major modes: Classroom Room Trainings, Online instructor Led Trainings, Self-paced e-learning trainings, and Corporate Trainings.

About Course?

This is an Artificial Intelligence Engineer Master Course that is a comprehensive learning approach for mastering the domains of Artificial Intelligence, Data Science, Business Analytics, Business Intelligence, Python coding, and Deep Learning with TensorFlow. Upon completion of the training, you will be able to take on challenging roles in the artificial intelligence domain.

why we take course?

Artificial intelligence is one of the hottest domains being heralded as the one with the ability to disrupt companies cutting across industry sectors. This Sadiq Tech Solutions Artificial Intelligence Engineer Master Course will equip you with all the necessary skills needed to take on challenging and exciting roles in the artificial intelligence, data science, business analytics, Python, R statistical computing domains and grab the best jobs in the industry at top-notch salaries.

Courses Covered in AI Diploma:

Course 1: As a Future ready IT employee I am interested to learn foundations of Data Analytics with statistics and Tableau, R

Course 2: I want to redefine my Analytical knowledge into python programming for Machine Learning Engineer

Course 3: I want to apply statistics and Python on Machine Learning models for Predictions& classifications of Data in various industry segments for intelligent Business

Course 4: So now I am ready to apply machine Learning models to implement Recommendation systems and Creating Machine Learning service in the cloud(IBM WATSON,AWS)

Course 5: Understanding tools of Bigdata and implementation using Machine Learning

Course 6: Applying Machine learning to speech Analytics to process Text using NLP

Course 7: Start Learning Neural Networks using Tensor flows and Keras for image classification and Data Extraction from Image(OCR)

Course 8: Now I am ready to face defection of human, Object using computer vision Technology with Open CV

Course 9: Sensing Real world Data and transforming it to Intelligent actions using IOT

Artificial intelligence vs Machine Learning vs Deep Learning

Nowadays many misconceptions are there related to the words **machine learning**, **deep learning** and **artificial intelligence(AI)**, most of the people think all these things are same whenever they hear the word AI, they directly relate that word to machine learning or vice versa, well yes, these things are related to each other but not the same. Let"s see how.

Machine Learning:

Before talking about machine learning lets talk about another concept that is called data mining. Data mining is a technique of examining a large pre-existing database and extracting new information from that database, it"s easy to understand, right, machine learning does the same, in fact, machine learning is a type of data mining technique.

Here"s is a basic definition of machine learning –

"Machine Learning is a technique of parsing data, learn from that data and then apply what they have learned to make an informed decision"

Now a day's many of big companies use machine learning to give there users a better experience, some of the examples are, Amazon using machine learning to give better product choice recommendations to there costumers based on their preferences, Netflix uses machine learning to give better suggestions to their users of the Tv series or movie or shows that they would like to watch.

Deep Learning:

Deep learning is actually a subset of machine learning. It technically is machine learning and functions in the same way but it has different capabilities.

The main difference between deep and machine learning is, machine learning models become better progressively but the model still needs some guidance. If a machine learning model returns an inaccurate prediction then the programmer needs to fix that problem explicitly but in the case of deep learning, the model does it by himself. Automatic car driving system is a good example of deep learning.

Let"s take an example to understand both machine learning and deep learning –

Suppose we have a flashlight and we teach a machine learning model that whenever someone says "dark" the flashlight should be on, now the machine learning model will analyse different phrases said by people and it will search for the word "dark" and as the word comes the flashlight will be on but what if someone said "I am not able to see anything the light is very dim", here the user wants the flashlight to be on but the sentence does not the consist the word "dark" so the flashlight will not be on. That"s where deep learning is different from machine learning. If it were a deep learning model it would on the flashlight, a deep learning model is able to learn from its own method of computing.

Artificial intelligence:

Now if we talk about AI, it is completely a different thing from Machine learning and deep learning, actually deep learning and machine learning both are the subsets of AI. There is no fixed definition for AI, you will find a different definition everywhere, but here is a definition that will give you idea of what exactly AI is.

"AI is a ability of computer program to function like a human brain "

AI means to actually replicate a human brain, the way a human brain thinks, works and functions. The truth is we are not able to establish a proper AI till now but we are very close to establish it, one of the examples of AI is Sophia, the most advanced AI model present today. The reason we are not able to establish proper AI till now is, we don"t know the many aspects of the human brain till now like why do we dream ? etc.

Why people relate machine learning and deep learning with artificial intelligence?

Machine learning and deep learning is a way of achieving AI, which means by the use of machine learning and deep learning we may able to achieve AI in future but it is not AI.

Machine Learning Introduction

Machine Learning is a field which is raised out of Artificial Intelligence(AI). Applying AI, we wanted to build better and intelligent machines. But except for few mere tasks such as finding the shortest path between point A and B, we were unable to program more complex and constantly evolving challenges. There was a realization that the only way to be able to achieve this task was to let machine learn from itself. This sounds similar to a child learning from its self. So machine learning was developed as a new capability for computers. And now machine learning is present in so many segments of technology, that we don"t even realize it while using it.

Finding patterns in data on planet earth is possible only for human brains. The data being very massive, the time taken to compute is increased, and this is where Machine Learning comes into action, to help people with large data in minimum time.

If big data and cloud computing are gaining importance for their contributions, machine learning as technology helps analyze those big chunks of data, easing the task of data scientists in an automated process and gaining equal importance and recognition.

The techniques we use for data mining have been around for many years, but they were not effective as they did not have the competitive power to run the algorithms. If you run deep learning with access to better data, the output we get will lead to dramatic breakthroughs which is machine learning.

Kinds of Machine Learning

There are three kinds of Machine Learning Algorithms.
- ✓ Supervised Learning
- ✓ Unsupervised Learning
- ✓ Reinforcement Learning

Supervised Learning

A majority of practical machine learning uses supervised learning.

In supervised learning, the system tries to learn from the previous examples that are given. (On the other hand, in unsupervised learning, the system attempts to find the patterns directly from the example given.)

Speaking mathematically, supervised learning is where you have both input variables (x) and output variables(Y) and can use an algorithm to derive the mapping function from the input to the output.

The mapping function is expressed as Y = f(X).

Example :
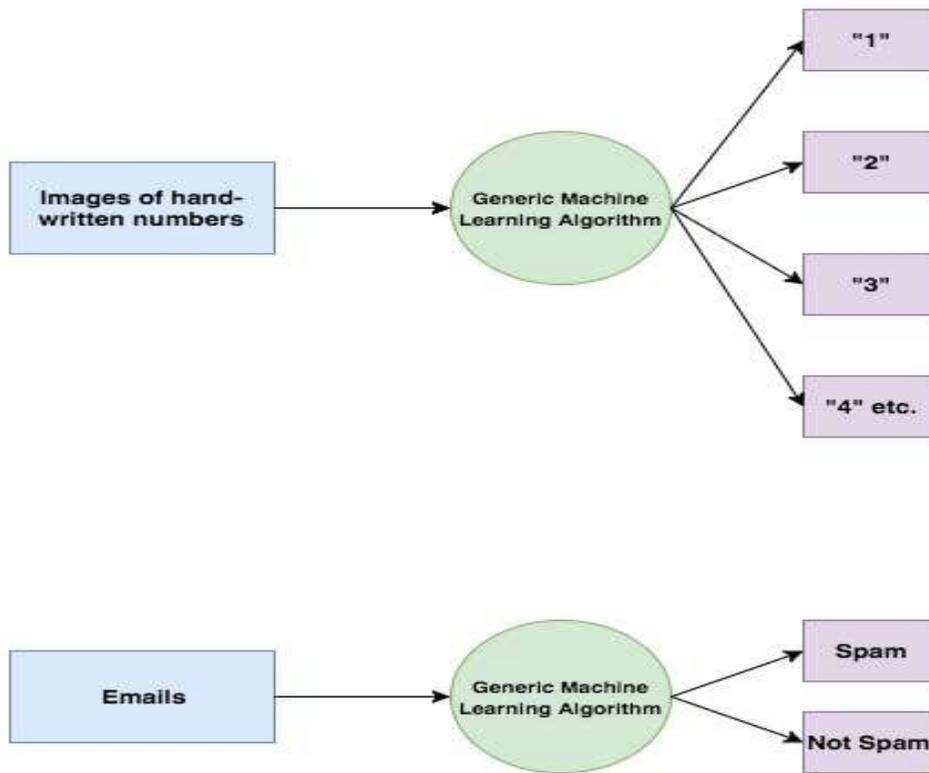
Supervised learning problems can be further divided into two parts, namely classification, and regression.

Classification: A classification problem is when the output variable is a category or a group, such as "black" or "white" or "spam" and "no spam".

Regression: A regression problem is when the output variable is a real value, such as "Rupees" or "height."

Unsupervised Learning:

In unsupervised learning, the algorithms are left to themselves to discover interesting structures in the data.

Mathematically, unsupervised learning is when you only have input data (X) and no corresponding output variables.

This is called unsupervised learning because unlike supervised learning above, there are no given correct answers and the machine itself finds the answers.

Unsupervised learning problems can be further divided into association and clustering problems.

Association: An association rule learning problem is where you want to discover rules that describe large portions of your data, such as "people that buy X also tend to buy Y".

Clustering: A clustering problem is where you want to discover the inherent groupings in the data, such as grouping customers by purchasing behavior.

Reinforcement Learning

A computer program will interact with a dynamic environment in which it must perform a particular goal (such as playing a game with an opponent or driving a car). The program is provided feedback in terms of rewards and punishments as it navigates its problem space.

Using this algorithm, the machine is trained to make specific decisions. It works this way: the machine is exposed to an environment where it continuously trains itself using trial and error method.
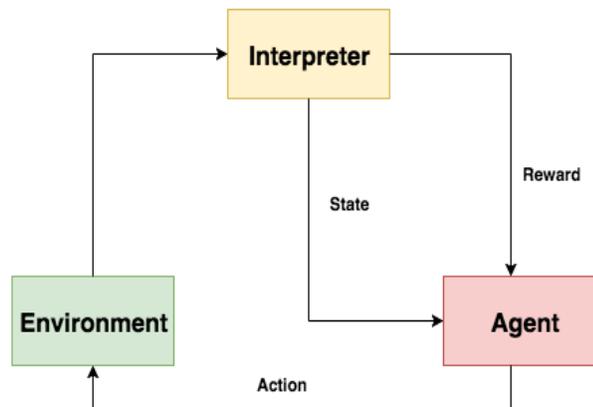
The Math of Intelligence

Machine Learning theory is a field that meets statistical, probabilistic, computer science and algorithmic aspects arising from learning iteratively from data which can be used to build intelligent applications.

Why Worry About The Maths?

There are various reasons why the mathematics of Machine Learning is necessary, and I will highlight some of them below:

Selecting the appropriate algorithm for the problem includes considerations of accuracy, training time, model complexity, the number of parameters and number of characteristics.

Identifying underfitting and overfitting by following the Bias-Variance tradeoff.

Choosing parameter settings and validation strategies.Estimating the right determination period and uncertainty.
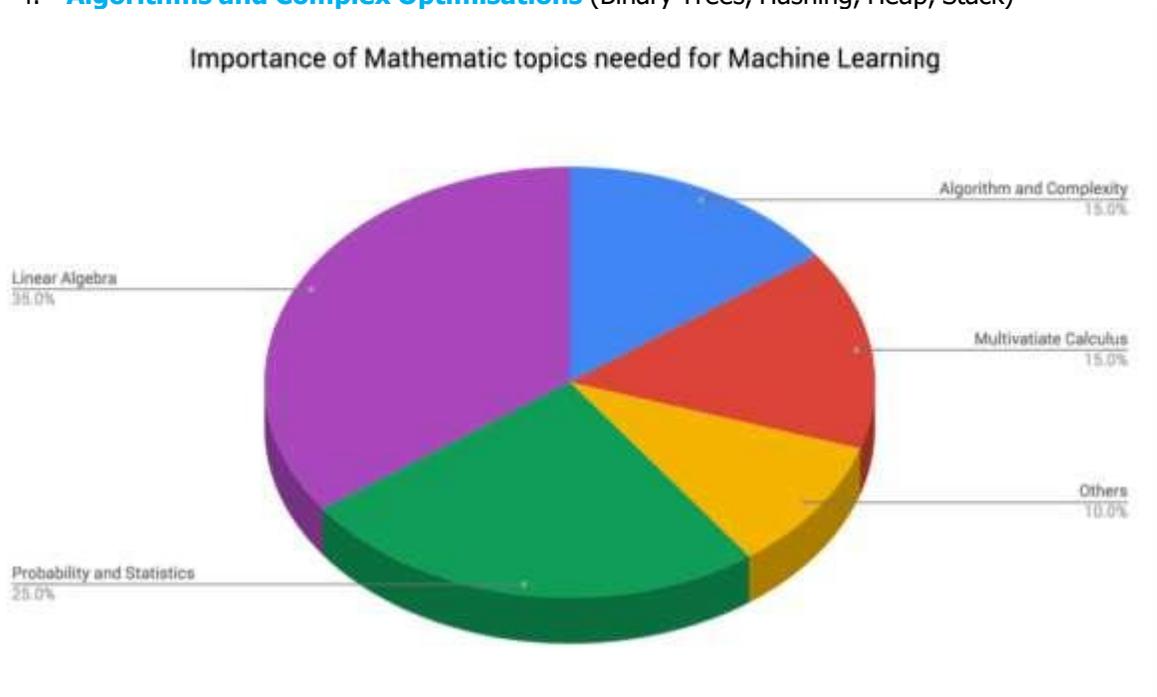
What Level of Maths Do We Need?

The foremost question when trying to understand a field such as Machine Learning is the amount of maths necessary and the complexity of maths required to understand these systems.

The answer to this question is multidimensional and depends on the level and interest of the individual.

Here is the minimum level of mathematics that is needed for Machine Learning Engineers / Data Scientists.

1. **Linear Algebra** (Matrix Operations, Projections, Factorisation, Symmetric Matrices, Orthogonalisation)

2. **Probability Theory and Statistics** (Probability Rules & Axioms, Bayes" Theorem, Random Variables, Variance and Expectation, Conditional and Joint Distributions, Standard Distributions.)

3. **Calculus** (Differential and Integral Calculus, Partial Derivatives)

4. **Algorithms and Complex Optimisations** (Binary Trees, Hashing, Heap, Stack)

Importance of Mathematic topics needed for Machine Learning

Algorithm and Complexity 15.0%

Multivatiate Calculus 15.0%

Others 10.0%

Probability and Statistics 25.0%

Linear Algebra 35.0%

Regression Models

✓ In this Machine Learning tutorial, we are going to get the basic understanding of what exactly the Regression Models in Machine Learning.

✓ Regression Models are the most popular among all statistical models which are generally used to estimate the relationship between variables.

✓ There are different types of Regression Models. Out of which below are the most commonly used models for Regression Analysis.

1. Simple Linear Regression
2. Multiple Linear Regression
3. Polynomial Regression
4. Logistic Regression

5. Ridge Regression
6. Lasso Regression
7. ElasticNet Regression
8. Support Vector Regression

These above models will be discussed elaborately in the next upcoming topics in this Machine learning tutorial.

Understanding Regression Analysis

Regression Analysis involves in creating the machine learning models which predict a numeric value. Prediction always happens with a solid machine learning model which estimates the relationship between a dependent variable and Independent variable.
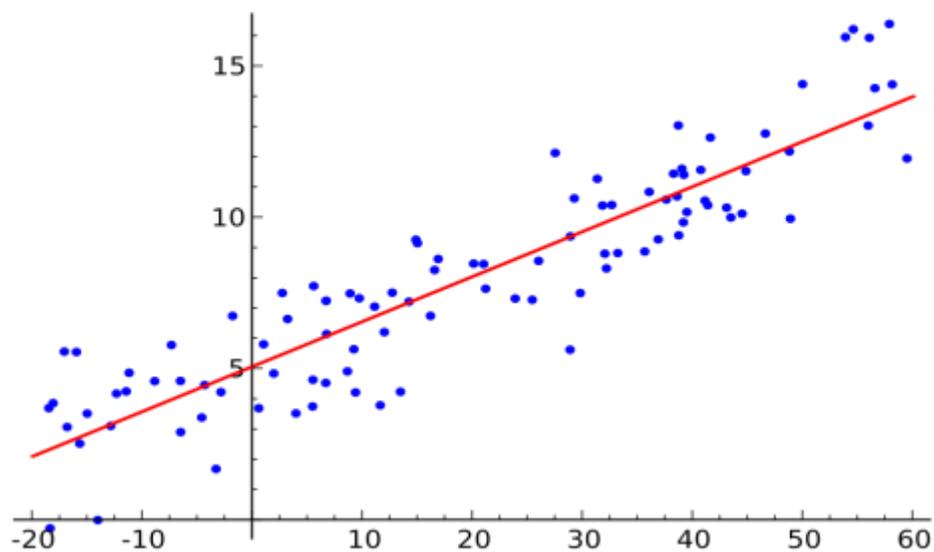
From the above graph, it can be understood the Independent or Predictor variable is on the X-axis,
Whereas the dependent or response variable is on the Y-axis. Coming to the inclined line is none other than the Regression Line. And the Plotted data points can be seen as blue colored dots.

What are Dependent and Independent Variables?

Dependent Variable: By the name itself we can clearly understand that this variable will vary depending on other variables or other factors.
Dependent variable is also called as Response variable (Outcome).

Example:

Consider the students score in the examination which could vary based on several factors.

Independent Variable: This is the variable which is not influenced by other variable rather we can say this variable standalone which can have a quality of influencing others.
Independent variable is also called as Predictor variable (Input).

Consider the same example as student score in the examination. Generally, student score will depend on various factors like hours of study, attendance etc. So, the time spent by the student to prepare for examination can be considered as independent variable.

Let us continue to learn the Simple Linear Regression....

Simple Linear Regression

Independent and Dependent variables:

In the context of Statistical learning, there are two types of data:
- ✓ **Independent variables**: Data that can be controlled directly.
- ✓ **Dependent variables**: Data that cannot be controlled directly.

The data that can"t be controlled i.e. dependent variables need to predicted or estimated.

Model:
A model is a transformation engine that helps us to express dependent variables as   a function of independent variables.
Parameters: Parameters are ingredients added to the model for estimating the   output.

Concept
Linear regression models provide a simple approach towards supervised learning. They are simple yet effective.

Wait, what do we mean by linear?

Linear implies the following: arranged in or extending along a straight or nearly straight line. Linear suggests that the relationship between dependent and independent variable can be **expressed in a straight line.**

Recall the geometry lesson from high school. What is the equation of a line?

y = mx + c

Linear regression is nothing but a manifestation of this simple equation.
- ✓ **y** is the dependent variable i.e. the variable that needs to be estimated and predicted.
- ✓ **x** is the independent variable i.e. the variable that is controllable. It is the input.
- ✓ **m** is the slope. It determines what will be the angle of the line. It is the parameter denoted as β.
- ✓ **c** is the intercept. A constant that determines the value of y when x is 0.

George Box, a famous British statistician, once quoted:
"All models are wrong; some are useful."

Linear regression models are not perfect. It tries to approximate the relationship between dependent and independent variables in a straight line. Approximation leads to errors. Some errors can be reduced. Some errors are inherent in the nature of the problem. These errors cannot be eliminated. They are called as an **irreducible error**, the noise term in the true relationship that cannot fundamentally be reduced by any model.

The same equation of a line can be re-written as:

$$Y = \beta_0 + \beta_1 X + \epsilon$$

β0 and β1 are two unknown constants that represent the intercept and slope. They are the parameters.

ε is the error term.

Formulation

Let us go through an example to explain the terms and workings of a Linear regression model.

Fernando is a Data Scientist. He wants to buy a car. He wants to estimate or predict the car price that he will have to pay. He has a friend at a car dealership company. He asks for prices for various other cars along with a few characteristics of the car. His friend provides him with some information.

make	fuelType	nDoors	engineSize	price
alfa-romero	gas	two	130	13495
alfa-romero	gas	two	130	16500
alfa-romero	gas	two	152	16500
audi	gas	four	109	13950
audi	gas	four	136	17450
audi	gas	two	136	15250
audi	gas	four	136	17710
audi	gas	four	136	18920
audi	gas	four	131	23875

The following are the data provided to him:
make: make of the car.

fuelType: type of fuel used by the car.
nDoor: number of doors.

engineSize: size of the engine of the car.

price: the price of the car.

First, Fernando wants to evaluate if indeed he can predict car price based on engine size. The first set of analysis seeks the answers to the following questions:
- ✓ Is price of car price related with engine size?
- ✓ How strong is the relationship?
- ✓ Is the relationship linear?
- ✓ Can we predict/estimate car price based on engine size?

Fernando does a correlation analysis. Correlation is a measure of how much the two variables are related. It is measured by a metric called as the **correlation coefficient**. Its value is between 0 and 1.

If the correlation coefficient is a large(> 0.7) +ve number, it implies that as one variable increases, the other variable increases as well. A large -ve number indicates that as one variable increases, the other variable decreases.

He does a correlation analysis. He plots the relationship between price and engine size.

He splits the data into training and test set. 75% of data is used for training. Remaining is used for the test.

He builds a linear regression model. He uses a statistical package to create the model. The model creates a linear equation that expresses **price of the car** as a function of **engine size.**
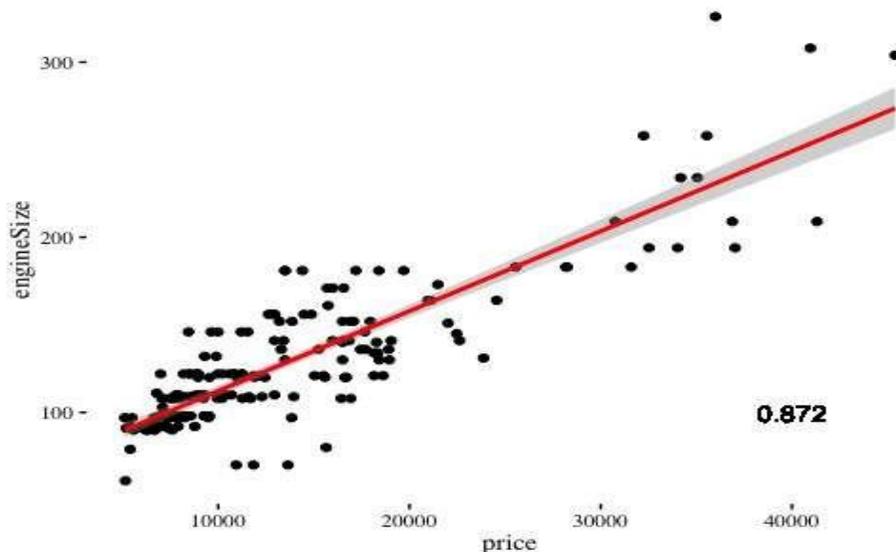
Following are the answers to the questions:
- ✓ Is price of car price related with engine size?

- ✓ Yes, there is a relationship.

- ✓ How strong is the relationship?

- ✓ The correlation coefficient is 0.872 => There is a strong relationship.

- ✓ Is the relationship linear?

- ✓ A straight line can fit => A decent prediction of price can be made using engine size.

- ✓ Can we predict/estimate the car price based on engine size?

- ✓ Yes, car price can be estimated based on engine size.

Fernando now wants to build a linear regression model that will estimate the price of the car price based on engine size. Superimposing the equation to the car price problem, Fernando formulates the following equation for price prediction.

price = β0 + β1 x engine size

Model Building and Interpretation

Model

Recall the earlier discussion, on how the data needs to be split into training and testing set. The training data is used to learn about the data. The training data is used to create the model. The testing data is used to evaluate the model performance.

Fernando splits the data into training and test set. 75% of data is used for training. Remaining is used for the test. He builds a linear regression model. He uses a statistical package to create the model. The model produces a linear equation that expresses **price of the car** as a function of **engine size.**
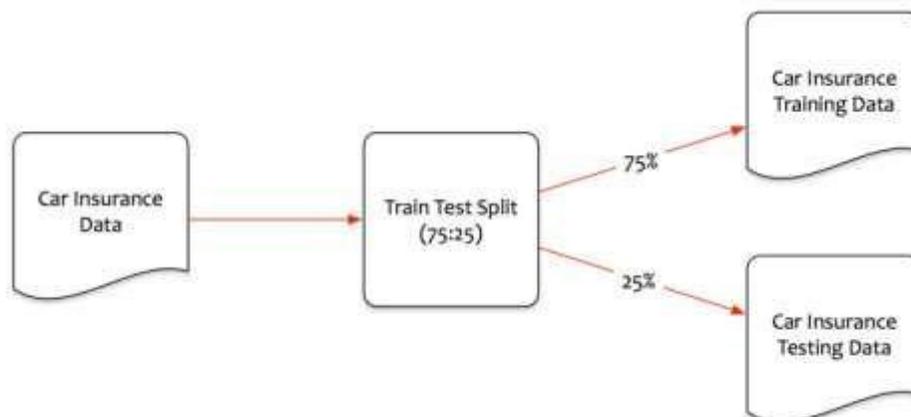
He splits the data into training and test set. 75% of data is used for training. Remaining is used for the test.

He builds a linear regression model. He uses a statistical package to create the model. The model creates a linear equation that expresses **price of the car** as a function of **engine size.**
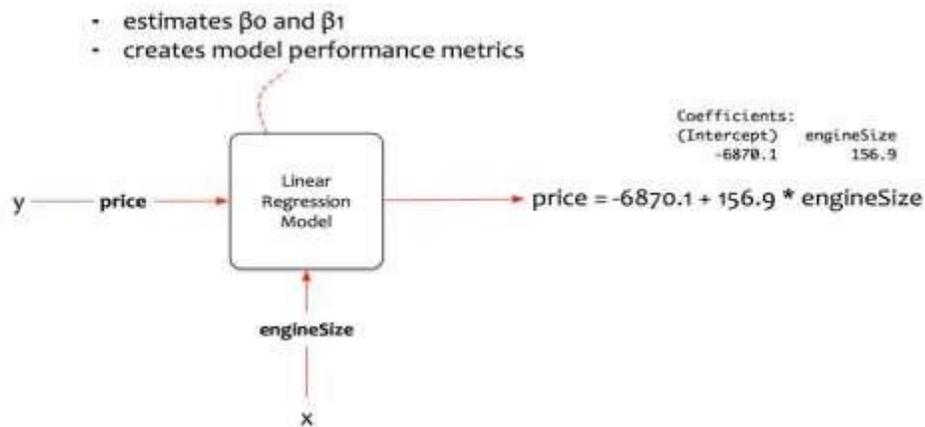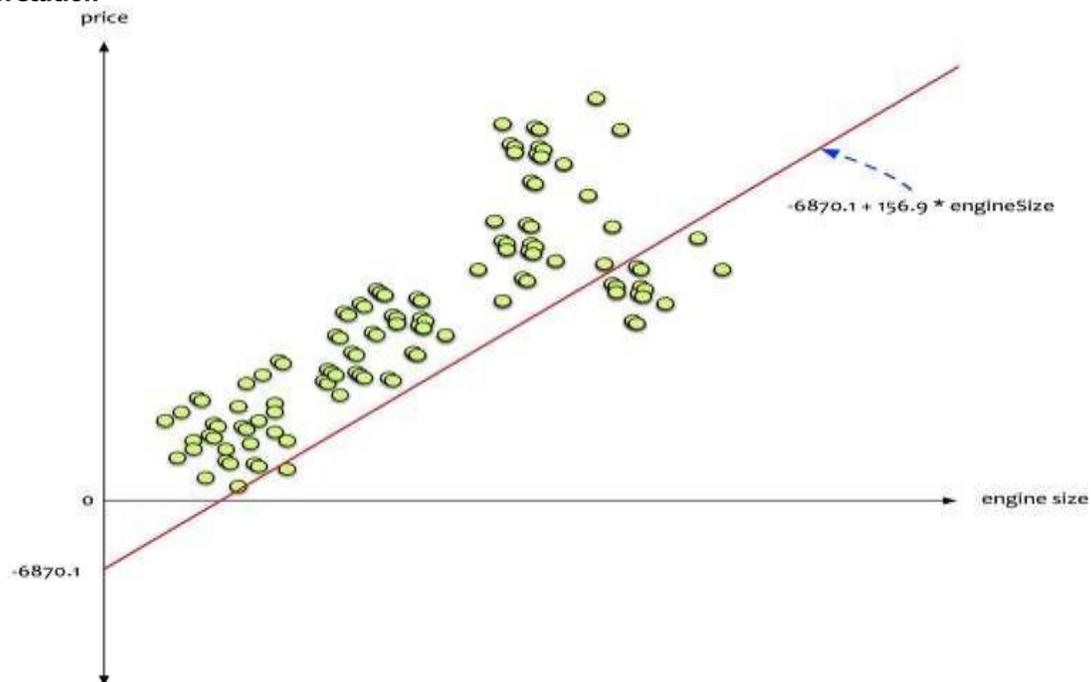

The model estimates the parameters:
- β0 is estimated as -6870.1

- β1 is estimated as 156.9

The linear equation is estimated as:

price = -6870.1 + 156.9 x engine size

Interpretation

Note:

- This is a preview of the Data Science with AI eBook containing only 15 pages.

- It is provided to help you understand how the full Data Science with AI eBook looks and is structured.

- The complete Data Science with AI eBook includes detailed concepts, real-world examples,
   and career guidance.

- Purchase the full Data Science with AI eBook for just ₹499.
  (Don't miss any offer is going on)

- Buy now from our official website: https://topitcourses.com/data-science-with-ai/