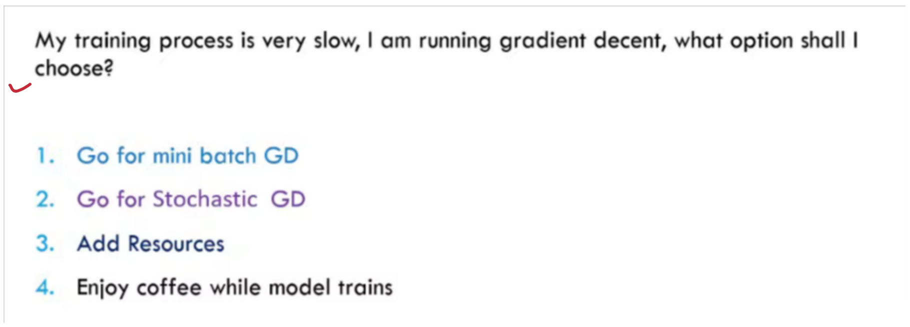
Session 8

Tuesday, April 11, 2023 6:34 PM



My training process is very slow, I am running gradient decent, what option shall I choose?

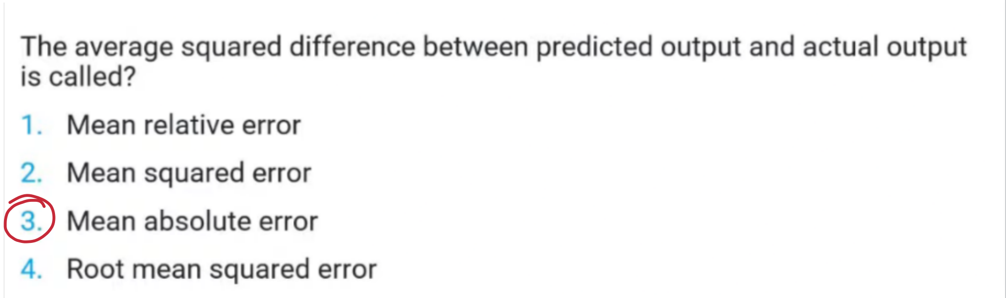
1. Go for mini batch GD
2. Go for Stochastic GD
3. Add Resources
4. Enjoy coffee while model trains

1. Mini-batch Gradient Descent (Mini-batch GD): Mini-batch GD is a variation of gradient descent where instead of computing the gradients and updating the parameters for the entire dataset, you use a smaller subset or "mini-batch" of the data at each iteration. This can speed up the training process as it allows you to update the parameters more frequently and take advantage of vectorized operations for improved efficiency.
2. Stochastic Gradient Descent (SGD): SGD is another variation of gradient descent where you update the parameters using only one example from the training dataset at a time, as opposed to using mini-batches or the entire dataset. While this can lead to more frequent updates and faster convergence, it can also introduce more randomness into the training process due to the variability of individual examples.

Choosing between mini-batch GD and SGD depends on your specific problem and dataset characteristics. Here are some factors to consider:

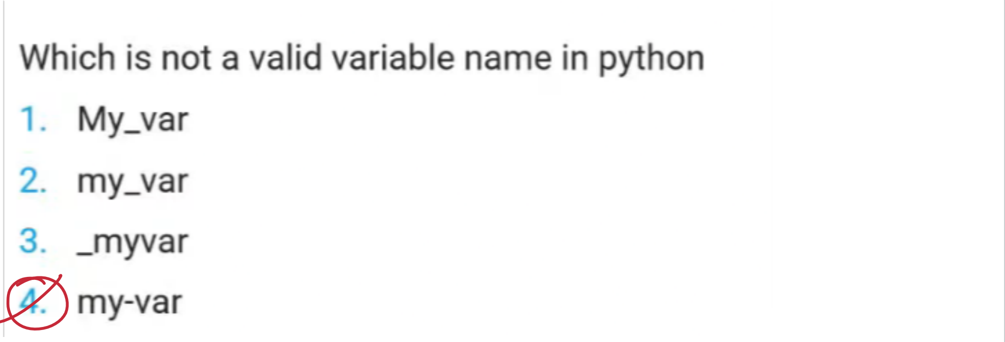
- Dataset size: If you have a large dataset, mini-batch GD may be a good option as it allows you to take advantage of parallelization and compute gradients and update parameters more frequently. If you have a smaller dataset, SGD may be suitable as it can converge faster with fewer updates.
- Computational resources: Mini-batch GD may require more computational resources, such as memory and processing power, as compared to SGD, as it operates on larger subsets of data. If you have limited computational resources, SGD may be a more feasible option.
- Convergence speed: Mini-batch GD typically converges faster than SGD as it uses a larger subset of data for updates. However, SGD can be more erratic in its convergence due to the randomness introduced by using individual examples. If you prioritize faster convergence, mini-batch GD may be a better option.
- Noise tolerance: SGD can introduce more randomness into the training process due to the use of individual examples, which can be beneficial in some cases as it allows the model to escape local optima. If your problem is noisy or has local optima, SGD may be more suitable.

Overall, both mini-batch GD and SGD have their pros and cons, and the best option for your specific case depends on your dataset, computational resources, convergence speed requirements, and noise tolerance. It's recommended to experiment with both methods and choose the one that works best for your particular situation.



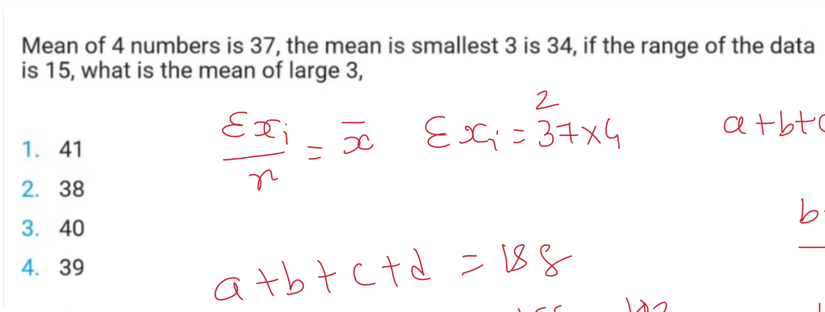
The average squared difference between predicted output and actual output is called?

1. Mean relative error
2. Mean squared error
3. [✓ Mean absolute error]
4. Root mean squared error



Which is not a valid variable name in python

1. My_var
2. my_var
3. _myvar
4. [✓ my-var]



Mean of 4 numbers is 37, the mean is smallest 3 is 34, if the range of the data is 15, what is the mean of large 3,

1. 41
2. 38
3. 40
4. 39

$\frac{\Sigma x_i}{n} = \bar{x}$ $\Sigma x_i = 37 \times 4$ (2)

$a+b+c+d = 148$

$a+b+c+d = 148$

$a = 158 - 102$

$a = 56$

$\frac{b+c+d}{3} = 34$

$b+c+d = (30+4)3$

$= 90+12$

$b+c+d = 102$

$a - d = 15$

$56 - d = 15$

$d = 41$

$b+c = 102 - 41$

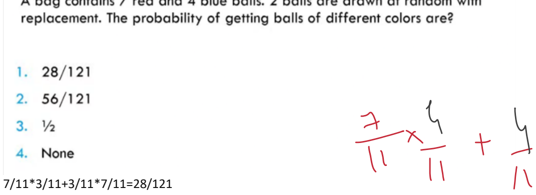
$b+c = 61$

$\frac{61+56}{3}$

$\frac{127}{3}$

$= 39$ $\boxed{39}$

A bag contains 7 red and 4 blue balls, 2 balls are drawn at random with

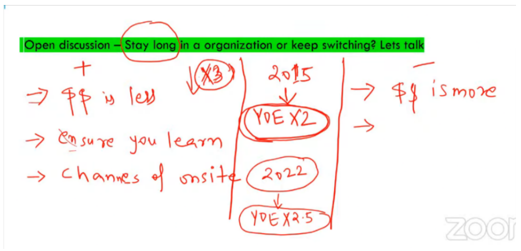


A bag contains 7 red and 4 blue balls. 2 balls are drawn at random with replacement. The probability of getting balls of different colors are?

1. 28/121
2. 56/121
3. ½
4. None

$$\frac{7}{11} \times \frac{4}{11} + \frac{4}{11} \times \frac{7}{11} = \frac{28}{121}$$

7/11*3/11+3/11*7/11=28/121



Open discussion – Stay long in a organization or keep switching? Lets talk

Stay: Scope of work, Salary Growth - 10-20% Yearly Growth
Good Salary = Exp in year * 2, before 2015 , Exp in year * 2.5, after 2015
Data Science = Exp in year * 3 , Min 30Lakh