**Artificial Intelligence Quarter 2 Batch 35 PIAIC\_91502**

Give definition of each main and sub topic (if any) in 3 lines.

1. keras gradients types

**Gradient Descent** is an optimization algorithm used for minimizing the cost function in various machine learning algorithms. It is basically used for updating the parameters of the learning model.

**Types of gradient Descent:**

1. **Batch Gradient Descent:**This is a type of gradient descent which processes all the training examples for each iteration of gradient descent. But if the number of training examples is large, then batch gradient descent is computationally very expensive.
2. **Stochastic Gradient Descent:** This is a type of gradient descent which processes 1 training example per iteration. Hence, the parameters are being updated even after one iteration in which only a single example has been processed. Hence this is quite faster than batch gradient descent.
3. **Mini Batch gradient descent:** Mini-batch gradient is a variation of stochastic gradient descent where instead of single training example, mini-batch of samples is used. It’s one of the most popular optimization algorithms. It works faster.

2)layer topology like inception block, multi head and one more

**inception block**

Inception Modules are used in [Convolutional Neural Networks](https://deepai.org/machine-learning-glossary-and-terms/convolutional-neural-network) to allow for more efficient computation and deeper Networks through a dimensionality reduction with stacked 1×1 convolutions. The modules were designed to solve the problem of computational expense, as well as overfitting, among other issues. The solution, in short, is to take multiple kernel filter sizes within the CNN, and rather than stacking them sequentially, ordering them to operate on the same level.

**Multiple head**

We can use the functional API to build models with multiple outputs (or multiple *heads*). A simple example is a network that attempts to simultaneously predict different properties of the data, such as a network that takes as input a series of social media posts from a single anonymous person and tries to predict attributes of that person, such as age, gender, and income level. Importantly, training such a model requires the ability to specify different loss functions for different heads of the network:The simplest way to combine different losses is to sum them all. In Keras, we can use either a list or a dictionary of losses in compile to specify different objects for different

outputs; the resulting loss values are summed into a global loss, which is minimized during training.