**Assignment\_5**

1. Why would you want to use the Data API?

**Ans: An API (Application Programming Interface) is a set of functions that allows applications to access data and interact with external software components, operating systems, or microservices. To simplify, an API delivers a user response to a system and sends the system's response back to a user.**

1. What are the benefits of splitting a large dataset into multiple files?

**Ans:**

* **Multiple Users can Access Data Simultaneously. When the data in a database is split, in frontend and backend, it can be easily supplied to multiple users sharing a network.**
* **Provides Better Protection.**
* **Allows for Future Planning.**
* **Easy to Modify User Interface.**

1. During training, how can you tell that your input pipeline is the bottleneck? What can you do to fix it?

**Ans: A CPU bottleneck occurs when the GPU resource is under utilized as a result of one, or more of the CPUs, having reached maximum utilization. In this situation, the GPU will be partially idle while it waits for the CPU to pass in training data. This is an undesired state.**

1. Can you save any binary data to a TFRecord file, or only serialized protocol buffers?

**Ans:**

**Step 1: Create a Free Roboflow Public Workspace. Roboflow is the universal conversion tool for computer vision annotation formats. ...**

**Step 2: Upload your data into Roboflow. ...**

**Step 3: Generate Dataset Version.**

**Step 4: Export Dataset Version**.

1. Why would you go through the hassle of converting all your data to the Example protobuf format? Why not use your own protobuf definition?

**Ans:** [**Protocol buffers**](https://developers.google.com/protocol-buffers/)**are a mechanism for sending data through the series of tubes known as the Internet. One common use of them is to define [gRPC](https://grpc.io/" \t "_blank) specifications — essentially a form of remote procedure calls. With gRPC service definitions, you create a “service” that has RPC methods. These RPC methods take a request “message” and return a response “message”.**

1. When using TFRecords, when would you want to activate compression? Why not do it systematically?

7.Data can be preprocessed directly when writing the data files, or within the tf.data pipeline, or in preprocessing layers within your model, or using TF Transform. Can you list a few pros and cons of each option?

**Ans:**

**Use the Keras preprocessing layers, such as tf. keras. layers. Resizing , tf. keras. layers. Rescaling , tf. keras. layers. ...**

**Use the tf. image methods, such as tf. image. flip\_left\_right , tf. image. rgb\_to\_grayscale , tf. image. adjust\_brightness , tf. image.**