✓ **Congratulations! You passed!**

Grade received 100%   To pass 80% or higher

Go to next item

Week 4 Quiz

Total points 9

1. What does the acronym ETL stand for?   `1 / 1 point`

   ○ External / Transform / Load

   ○ Extract / Transfer / Load

   ⦿ Extract / Transform / Load

   ○ Enhance / Transfer / Load

   ✓ **Correct**

2. You have a multi processor machine, containing a CPU and GPU. How would you typically distribute these for training a model?   `1 / 1 point`

   ○ Use CPU and GPU for all tasks in parallel

   ○ Use CPU for extracting and the GPU for transforming and loading

   ○ Use CPU for extracting and loading, and the GPU for transforming

   ⦿ Use CPU for extracting, transferring and loading, and the GPU for training

   ✓ **Correct**

3. One way to speed up ETL is to use a cache. What's the API for this called?   `1 / 1 point`

   ○ tf.data.DataCache()

   ○ tf.data.Dataset.ETLCache()

   ⦿ tf.data.Dataset.cache()

   ○ tf.data.Dataset.datacache()

   ✓ **Correct**

4. I have a dataset loaded using this code: `1 / 1 point`

```
1   dataset = tfds.load('cats_vs_dogs',split=tfds.Split.TRAIN)
```

   How would I cache it on disk?

○ train_dataset = dataset.cache(filename='cache')

○ train_dataset = dataset.cache()

○ train_dataset = dataset.cache(file='cache')

○ train_dataset = dataset.cache(cachename=file)

✓ **Correct**

---

5. I have a dataset loaded using this code: **1 / 1 point**

```
1    dataset = tfds.load('cats_vs_dogs',split=tfds.Split.TRAIN)
```

How would I cache it in memory?

○ train_dataset = dataset.memorycache()

○ train_dataset = dataset.cache_in_memory()

◉ train_dataset = dataset.cache()

○ train_dataset = dataset.cache(cachename='memory')

✓ **Correct**

---

6. If I create a function called 'augment' that transforms data, what code would I use to apply this after loading a **1 / 1 point**
dataset with

```
1    dataset = tfds.load('cats_vs_dogs',split=tfds.Split.TRAIN)
```

○ augmented_dataset = map(augment)

○ augmented_dataset = dataset.augment(dataset)

◉ augmented_dataset = dataset.map(augment)

○ augmented_dataset = dataset.augment()

✓ **Correct**

---

7. If you want to parallelise the transform of a dataset across multiple cores, what's the correct call? **1 / 1 point**

○ s = dataset.map(augment, num_parallel=2)

◉ s = dataset.map(augment, num_parallel_calls=2)

○ s = dataset.map(augment, 2)

○ s = dataset.map(augment, parallel_calls=2)

✓ **Correct**

8. If you're not sure how many cores are accessible, for example, if you're running in a shared cloud environment, how can you find out how many are available to you?

1 / 1 point

○ num_cores = multiprocessing.cpu.count()

○ It's not possible

○ num_cores = multiprocessing.available_cpus()

⦿ num_cores = multiprocessing.cpu_count()

⊘ **Correct**

9. The process of executing a custom map function over a batch of inputs is called:

1 / 1 point

○ Map batching

○ Batch mapping

○ Visualization

⦿ Vectorization

⊘ **Correct**