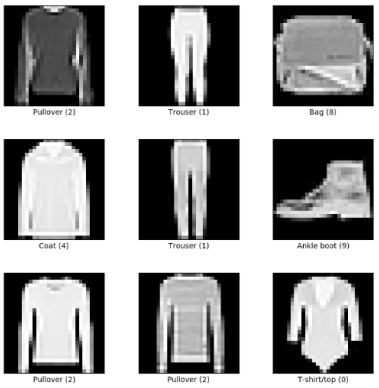
Differential Privacy Applied in Deep Learning

最後一組
M11215032 葉品和 M11215052 陳奕帆 M11215066 鄭宜珊

Dataset : fashion_mnist from keras

We use the “fashion_mnist” dataset from “keras”. There are a total of 60000 examples of training data and 10000 examples of testing data. The label has 10 different classes.



1.How do you perturb the data?

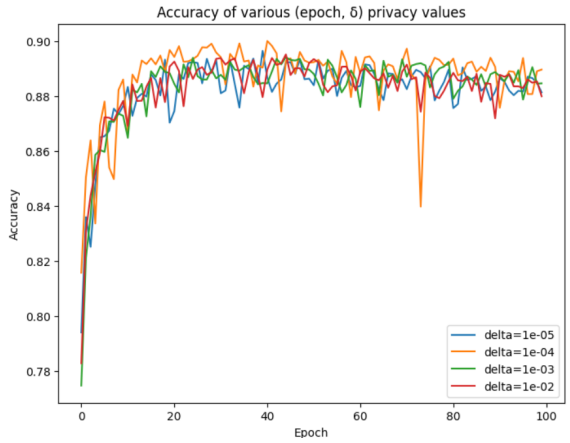
```
for idx in range(self._num_microbatches):
    # compute gradient
    microbatch_loss = tf.reduce_mean(tf.gather(microbatches_losses,  indices: [idx]))
    grads = gradient_tape.gradient(microbatch_loss, var_list)
    # accountant
    sample_state = self._dp_sum_query.accumulate_record(sample_params, sample_state, grads)

# add noise
grad_sums, self._global_state = (self._dp_sum_query.get_noised_result(sample_state, self._global_state))
```

1) define microbatch_loss and then compute gradient of every microbatch by using gradient_tape.gradient ()
2) update the set of previous microbatch gradients with the addition of the record argument
3) add noise to the gradient by using get_noised_result ()

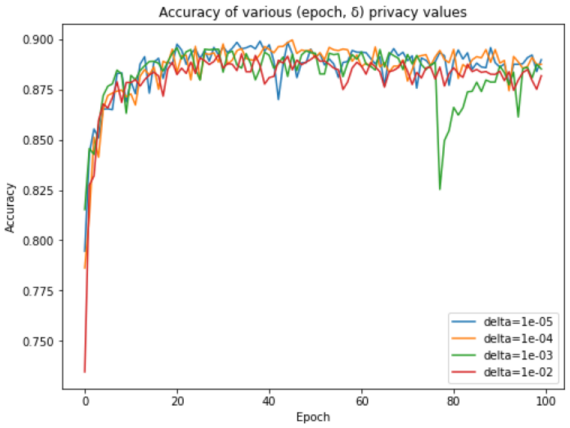
2.Effectiveness measure: Accuracy

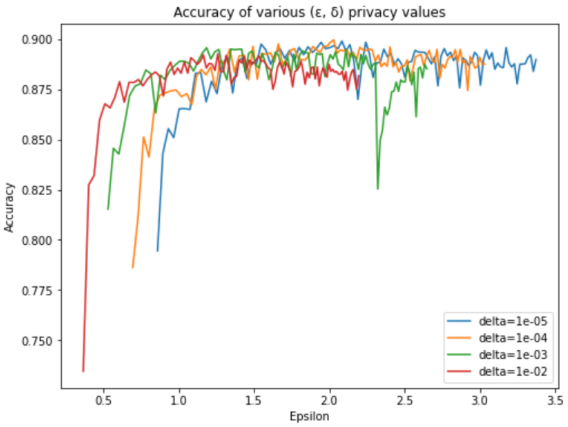
- learning rate = 0.001
- epoch = 100
- C = 1.0

σ (noise multiplier) = 1.1



σ (noise multiplier) = 1.2



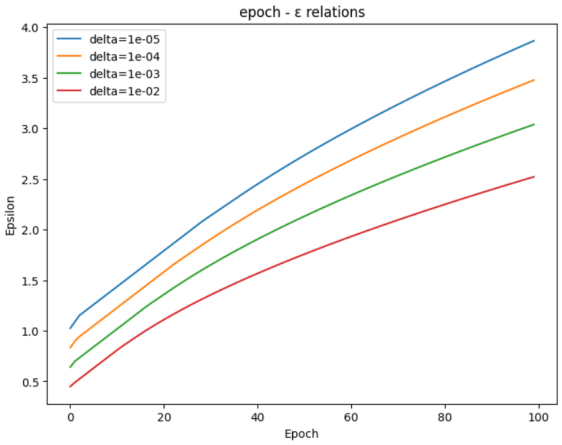


Since the y-axis in both graphs represents accuracy, with one x-axis being epoch and the other epsilon, we can observe the trend of accuracy increasing as epoch grows in both cases. By examining the epsilon graph simultaneously, we notice that epsilon also varies during the training process. This could be attributed to the update of model parameters or adjustments in the differential privacy mechanism.

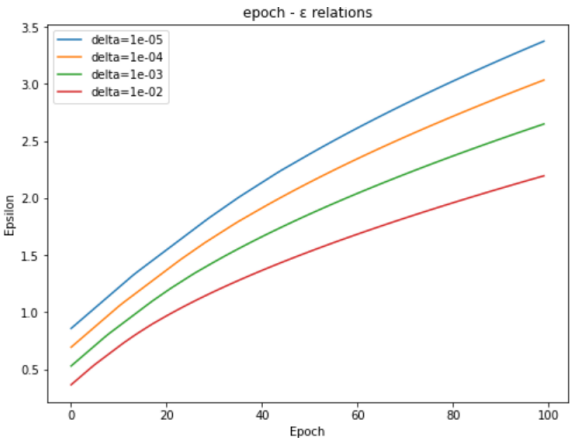
3.Privacy level: the value of ε

- learning rate = 0.001
- epoch = 100
- C = 1.0

σ (noise multiplier) = 1.1



σ (noise multiplier) = 1.2



A larger σ makes the noise curve flatter. The flatter curve means individual data points become less distinguishable because of the added noise. This leads to a flatter data distribution. The flatter distribution increases values deviating from the mean and makes pinpointing specific data entries harder. As a result, privacy protection strengthens as reflected in the lower ε.