

## Section 4: Attention

Now that you are done with the videos of section 1, let's assess your learning. Here, are a few questions, followed by 4 options, out of which 1 is the correct option. Select the right option and validate your learning! The answers are provided in a separate sheet

Q1. What layer does Attention use to learn weight probabilities

- a. Softmax
- b. Cosine
- c. Sigma

Q2. Attention can be computed on single strings

- a. True
- b. False

Q3. This layer type can apply a model multiple times to an input

- a. TimeDistributed
- b. Loop
- c. Map

Q4. Keras models can include whole Keras models as layers

- a. True
- b. False

Q5. Networks of networks can be combined to make predictions and are called

- a. Ensembles
- b. Aggregates

Q6. Ensemble networks have smaller numbers of parameters than single networks

- a. True
- b. False