

1. **(1 pt.)** Name the first prominent neural network term embedding technique published by Mikolov et al. in 2013?
2. **(1 pts.)** Match the component description on the left with the architecture on the right. Each architecture can have more than one match.

Inputs one token at a time.

Can process inputs in forward or backward directions.

Recurrent Neural Network

Relies on attention to highlight important information in token sequences.

Relies on recurrent units to preserve information across token sequences.

Transformer

Inputs all tokens at once.

3. **(1 pt.) Circle the correct phrase: (In-context learning / Retrieval augmented generation)** augments prompts with examples and information necessary to address a specific user query. **(In-context learning / Retrieval augmented generation)** augments prompts with relevant search results to ground query responses.

Use [HuggingFace.co](https://huggingface.co) and locate the [H2O.ai](https://huggingface.co/H2Oai) Danube 3 500M Chat model to answer Questions 4—7.

4. **(1 pt.)** What is the total number of parameters?
5. **(1 pt.)** Roughly how many Gigabytes (GB) of RAM are required to work with this model?
6. **(1 pt.)** How many transformer blocks does the model include?
7. **(1 pt.)** Please specify the architecture of the last block exactly. (Copy and paste from [HuggingFace.co](https://huggingface.co) information.)

Use the assignment notebook to complete Questions 8—10. Before running the notebook, copy it to your own drive and follow these instructions carefully:

- In cell 4, specify the correct maximum length of text inputs.
 - In cell 5, specify:
 - The correct input dimension for the fixed-width embedding model.
 - The correct number of hidden units and activation functions in the “x2d” layer.
(HINT: Look at the name of the layer and think about the fact that its output will be used for plotting below.)
 - In cell 6, select an appropriate number of epochs for training to achieve a high accuracy.
 - In cell 7, specify the correct data for making predictions.
 - In cell 9, specify the correct embedding model layer for extracting term embeddings.
8. **(1 pt.)** What is the total number of weights in the embedding model? What percentage are trainable?
9. **(1 pt.)** What are the dimensions of the extracted features?
10. **(1 pt.)** Paste your plot of term embeddings (output from cell 10) with an over-layed boundary line between positive and negative words here.