1. Implement the Continuous Bag of Words (CBOW) Model. Stages can be:  
   a. Data preparation  
   b. Generate training data  
   c. Train model  
   d. Output

In [1]:

**import** tensorflow **as** tf

**from** tensorflow.keras.models **import** Sequential

**from** tensorflow.keras.layers **import** Dense,\

Embedding, Lambda

**from** tensorflow.keras.preprocessing.text **import** Tokenizer

**import** numpy **as** np

**import** matplotlib.pyplot **as** plt

**from** sklearn.decomposition **import** PCA

**import** re

2023-11-05 10:46:30.013620: I tensorflow/core/platform/cpu\_feature\_guard.cc:182] This TensorFlow binary is optimized to use available CPU instructions in performance-critical operations.

To enable the following instructions: AVX2 FMA, in other operations, rebuild TensorFlow with the appropriate compiler flags.

**a. Data preparation**

In [2]:

data **=** """We are about to study the idea of a computational process.

Computational processes are abstract beings that inhabit computers.

As they evolve, processes manipulate other abstract things called data.

The evolution of a process is directed by a pattern of rules

called a program. People create programs to direct processes. In effect,

we conjure the spirits of the computer with our spells."""

In [3]:

*# for importing data from txt file*

*# with open("data.txt", "r", encoding="utf-8") as file:*

*# data = file.read()*

In [4]:

sentences **=** data**.**split(".")

In [5]:

sentences

Out[5]:

['We are about to study the idea of a computational process',

'\nComputational processes are abstract beings that inhabit computers',

'\nAs they evolve, processes manipulate other abstract things called data',

'\nThe evolution of a process is directed by a pattern of rules\ncalled a program',

' People create programs to direct processes',

' In effect,\nwe conjure the spirits of the computer with our spells',

'']

In [6]:

*#Clean Data*

clean\_sentences **=** []

**for** sentence **in** sentences:

*# skip empty string*

**if** sentence **==** "":

**continue**;

*# remove special characters*

sentence **=** re**.**sub('[^A-Za-z0-9]+', ' ', sentence)

*# remove 1 letter words*

sentence **=** re**.**sub(r'(?:^| )\w(?:$| )', ' ', sentence)**.**strip()

*# lower all characters*

sentence **=** sentence**.**lower()

clean\_sentences**.**append(sentence)

In [7]:

clean\_sentences

Out[7]:

['we are about to study the idea of computational process',

'computational processes are abstract beings that inhabit computers',

'as they evolve processes manipulate other abstract things called data',

'the evolution of process is directed by pattern of rules called program',

'people create programs to direct processes',

'in effect we conjure the spirits of the computer with our spells']

In [8]:

*# Define the corpus*

corpus **=** clean\_sentences

In [9]:

*# Convert the corpus to a sequence of integers*

tokenizer **=** Tokenizer()

tokenizer**.**fit\_on\_texts(corpus)

sequences **=** tokenizer**.**texts\_to\_sequences(corpus)

print("After converting our words in the corpus \

into vector of integers:")

print(sequences)

After converting our words in the corpus into vector of integers:

[[4, 5, 11, 6, 12, 1, 13, 2, 7, 8], [7, 3, 5, 9, 14, 15, 16, 17], [18, 19, 20, 3, 21, 22, 9, 23, 10, 24], [1, 25, 2, 8, 26, 27, 28, 29, 2, 30, 10, 31], [32, 33, 34, 6, 35, 3], [36, 37, 4, 38, 1, 39, 2, 1, 40, 41, 42, 43]]

In [10]:

*# creating dictionary for word to index and index to word*

index\_to\_word\_map **=** {}

word\_to\_index\_map **=** {}

**for** index\_1, sequence **in** enumerate(sequences):

print(sequence)

words\_in\_sentence **=** clean\_sentences[index\_1]**.**split()

print(words\_in\_sentence)

**for** index\_2, value **in** enumerate(sequence):

index\_to\_word\_map[value] **=** words\_in\_sentence[index\_2]

word\_to\_index\_map[words\_in\_sentence[index\_2]] **=** value

[4, 5, 11, 6, 12, 1, 13, 2, 7, 8]

['we', 'are', 'about', 'to', 'study', 'the', 'idea', 'of', 'computational', 'process']

[7, 3, 5, 9, 14, 15, 16, 17]

['computational', 'processes', 'are', 'abstract', 'beings', 'that', 'inhabit', 'computers']

[18, 19, 20, 3, 21, 22, 9, 23, 10, 24]

['as', 'they', 'evolve', 'processes', 'manipulate', 'other', 'abstract', 'things', 'called', 'data']

[1, 25, 2, 8, 26, 27, 28, 29, 2, 30, 10, 31]

['the', 'evolution', 'of', 'process', 'is', 'directed', 'by', 'pattern', 'of', 'rules', 'called', 'program']

[32, 33, 34, 6, 35, 3]

['people', 'create', 'programs', 'to', 'direct', 'processes']

[36, 37, 4, 38, 1, 39, 2, 1, 40, 41, 42, 43]

['in', 'effect', 'we', 'conjure', 'the', 'spirits', 'of', 'the', 'computer', 'with', 'our', 'spells']

In [11]:

print(index\_to\_word\_map)

print("\n")

print(word\_to\_index\_map)

{4: 'we', 5: 'are', 11: 'about', 6: 'to', 12: 'study', 1: 'the', 13: 'idea', 2: 'of', 7: 'computational', 8: 'process', 3: 'processes', 9: 'abstract', 14: 'beings', 15: 'that', 16: 'inhabit', 17: 'computers', 18: 'as', 19: 'they', 20: 'evolve', 21: 'manipulate', 22: 'other', 23: 'things', 10: 'called', 24: 'data', 25: 'evolution', 26: 'is', 27: 'directed', 28: 'by', 29: 'pattern', 30: 'rules', 31: 'program', 32: 'people', 33: 'create', 34: 'programs', 35: 'direct', 36: 'in', 37: 'effect', 38: 'conjure', 39: 'spirits', 40: 'computer', 41: 'with', 42: 'our', 43: 'spells'}

{'we': 4, 'are': 5, 'about': 11, 'to': 6, 'study': 12, 'the': 1, 'idea': 13, 'of': 2, 'computational': 7, 'process': 8, 'processes': 3, 'abstract': 9, 'beings': 14, 'that': 15, 'inhabit': 16, 'computers': 17, 'as': 18, 'they': 19, 'evolve': 20, 'manipulate': 21, 'other': 22, 'things': 23, 'called': 10, 'data': 24, 'evolution': 25, 'is': 26, 'directed': 27, 'by': 28, 'pattern': 29, 'rules': 30, 'program': 31, 'people': 32, 'create': 33, 'programs': 34, 'direct': 35, 'in': 36, 'effect': 37, 'conjure': 38, 'spirits': 39, 'computer': 40, 'with': 41, 'our': 42, 'spells': 43}

**b. Generate training data**

In [12]:

*# Define the parameters*

vocab\_size **=** len(tokenizer**.**word\_index) **+** 1

embedding\_size **=** 10

window\_size **=** 2

*# Generate the context-target pairs*

contexts **=** []

targets **=** []

**for** sequence **in** sequences:

**for** i **in** range(window\_size, len(sequence) **-** window\_size):

context **=** sequence[i **-** window\_size:i] **+** sequence[i **+** 1:i **+** window\_size **+** 1]

target **=** sequence[i]

contexts**.**append(context)

targets**.**append(target)

In [13]:

*# sample of training data*

**for** i **in** range(5):

words **=** []

target **=** index\_to\_word\_map**.**get(targets[i])

**for** j **in** contexts[i]:

words**.**append(index\_to\_word\_map**.**get(j))

print(words, "=>", target)

['we', 'are', 'to', 'study'] => about

['are', 'about', 'study', 'the'] => to

['about', 'to', 'the', 'idea'] => study

['to', 'study', 'idea', 'of'] => the

['study', 'the', 'of', 'computational'] => idea

In [14]:

*# Convert the contexts and targets to numpy arrays*

X **=** np**.**array(contexts)

Y **=** np**.**array(targets)

**c. Train model**

In [15]:

*# Define the CBOW model*

model **=** Sequential()

model**.**add(Embedding(input\_dim**=**vocab\_size, output\_dim**=**embedding\_size, input\_length**=**2 **\*** window\_size))

model**.**add(Lambda(**lambda** x: tf**.**reduce\_mean(x, axis**=**1)))

model**.**add(Dense(256, activation**=**'relu'))

model**.**add(Dense(512, activation**=**'relu'))

model**.**add(Dense(units**=**vocab\_size, activation**=**'softmax'))

*# Compile the model*

model**.**compile(loss**=**'sparse\_categorical\_crossentropy', optimizer**=**'adam', metrics**=**['accuracy'])

*# Train the model*

model**.**fit(X, Y, epochs**=**200, verbose**=**1)

Epoch 1/200

2/2 [==============================] - 0s 8ms/step - loss: 3.7840 - accuracy: 0.0294

Epoch 2/200

2/2 [==============================] - 0s 6ms/step - loss: 3.7747 - accuracy: 0.1765

Epoch 3/200

2/2 [==============================] - 0s 6ms/step - loss: 3.7670 - accuracy: 0.1176

Epoch 4/200

2/2 [==============================] - 0s 7ms/step - loss: 3.7590 - accuracy: 0.1176

Epoch 5/200

2/2 [==============================] - 0s 6ms/step - loss: 3.7490 - accuracy: 0.1176

Epoch 6/200

2/2 [==============================] - 0s 6ms/step - loss: 3.7379 - accuracy: 0.1176

Epoch 7/200

2/2 [==============================] - 0s 7ms/step - loss: 3.7251 - accuracy: 0.1176

Epoch 8/200

2/2 [==============================] - 0s 6ms/step - loss: 3.7099 - accuracy: 0.1176

Epoch 9/200

2/2 [==============================] - 0s 7ms/step - loss: 3.6919 - accuracy: 0.1176

Epoch 10/200

2/2 [==============================] - 0s 5ms/step - loss: 3.6711 - accuracy: 0.1471

Epoch 11/200

2/2 [==============================] - 0s 6ms/step - loss: 3.6471 - accuracy: 0.1471

Epoch 12/200

2/2 [==============================] - 0s 5ms/step - loss: 3.6184 - accuracy: 0.1176

Epoch 13/200

2/2 [==============================] - 0s 6ms/step - loss: 3.5851 - accuracy: 0.1176

Epoch 14/200

2/2 [==============================] - 0s 5ms/step - loss: 3.5477 - accuracy: 0.1176

Epoch 15/200

2/2 [==============================] - 0s 5ms/step - loss: 3.5061 - accuracy: 0.1176

Epoch 16/200

2/2 [==============================] - 0s 6ms/step - loss: 3.4633 - accuracy: 0.1176

Epoch 17/200

2/2 [==============================] - 0s 8ms/step - loss: 3.4173 - accuracy: 0.1176

Epoch 18/200

2/2 [==============================] - 0s 6ms/step - loss: 3.3707 - accuracy: 0.1176

Epoch 19/200

2/2 [==============================] - 0s 5ms/step - loss: 3.3228 - accuracy: 0.1176

Epoch 20/200

2/2 [==============================] - 0s 5ms/step - loss: 3.2718 - accuracy: 0.1176

Epoch 21/200

2/2 [==============================] - 0s 6ms/step - loss: 3.2240 - accuracy: 0.1176

Epoch 22/200

2/2 [==============================] - 0s 5ms/step - loss: 3.1789 - accuracy: 0.1176

Epoch 23/200

2/2 [==============================] - 0s 5ms/step - loss: 3.1368 - accuracy: 0.1471

Epoch 24/200

2/2 [==============================] - 0s 4ms/step - loss: 3.1041 - accuracy: 0.1471

Epoch 25/200

2/2 [==============================] - 0s 5ms/step - loss: 3.0692 - accuracy: 0.1471

Epoch 26/200

2/2 [==============================] - 0s 5ms/step - loss: 3.0417 - accuracy: 0.1471

Epoch 27/200

2/2 [==============================] - 0s 6ms/step - loss: 3.0187 - accuracy: 0.1471

Epoch 28/200

2/2 [==============================] - 0s 5ms/step - loss: 3.0006 - accuracy: 0.1176

Epoch 29/200

2/2 [==============================] - 0s 6ms/step - loss: 2.9701 - accuracy: 0.1176

Epoch 30/200

2/2 [==============================] - 0s 5ms/step - loss: 2.9323 - accuracy: 0.1176

Epoch 31/200

2/2 [==============================] - 0s 5ms/step - loss: 2.8934 - accuracy: 0.1176

Epoch 32/200

2/2 [==============================] - 0s 5ms/step - loss: 2.8603 - accuracy: 0.1471

Epoch 33/200

2/2 [==============================] - 0s 6ms/step - loss: 2.8345 - accuracy: 0.1765

Epoch 34/200

2/2 [==============================] - 0s 5ms/step - loss: 2.8075 - accuracy: 0.2353

Epoch 35/200

2/2 [==============================] - 0s 5ms/step - loss: 2.7839 - accuracy: 0.2647

Epoch 36/200

2/2 [==============================] - 0s 6ms/step - loss: 2.7592 - accuracy: 0.2647

Epoch 37/200

2/2 [==============================] - 0s 5ms/step - loss: 2.7336 - accuracy: 0.2647

Epoch 38/200

2/2 [==============================] - 0s 5ms/step - loss: 2.7042 - accuracy: 0.2353

Epoch 39/200

2/2 [==============================] - 0s 6ms/step - loss: 2.6726 - accuracy: 0.2353

Epoch 40/200

2/2 [==============================] - 0s 5ms/step - loss: 2.6350 - accuracy: 0.2647

Epoch 41/200

2/2 [==============================] - 0s 5ms/step - loss: 2.5918 - accuracy: 0.2353

Epoch 42/200

2/2 [==============================] - 0s 6ms/step - loss: 2.5551 - accuracy: 0.2647

Epoch 43/200

2/2 [==============================] - 0s 5ms/step - loss: 2.5199 - accuracy: 0.2647

Epoch 44/200

2/2 [==============================] - 0s 5ms/step - loss: 2.4888 - accuracy: 0.2353

Epoch 45/200

2/2 [==============================] - 0s 5ms/step - loss: 2.4622 - accuracy: 0.2353

Epoch 46/200

2/2 [==============================] - 0s 5ms/step - loss: 2.4347 - accuracy: 0.2647

Epoch 47/200

2/2 [==============================] - 0s 5ms/step - loss: 2.4077 - accuracy: 0.2941

Epoch 48/200

2/2 [==============================] - 0s 4ms/step - loss: 2.3822 - accuracy: 0.3235

Epoch 49/200

2/2 [==============================] - 0s 5ms/step - loss: 2.3580 - accuracy: 0.3235

Epoch 50/200

2/2 [==============================] - 0s 5ms/step - loss: 2.3299 - accuracy: 0.2941

Epoch 51/200

2/2 [==============================] - 0s 6ms/step - loss: 2.3047 - accuracy: 0.3235

Epoch 52/200

2/2 [==============================] - 0s 5ms/step - loss: 2.2747 - accuracy: 0.3235

Epoch 53/200

2/2 [==============================] - 0s 5ms/step - loss: 2.2469 - accuracy: 0.3235

Epoch 54/200

2/2 [==============================] - 0s 5ms/step - loss: 2.2179 - accuracy: 0.3235

Epoch 55/200

2/2 [==============================] - 0s 4ms/step - loss: 2.1988 - accuracy: 0.3529

Epoch 56/200

2/2 [==============================] - 0s 4ms/step - loss: 2.1861 - accuracy: 0.3824

Epoch 57/200

2/2 [==============================] - 0s 5ms/step - loss: 2.1612 - accuracy: 0.3824

Epoch 58/200

2/2 [==============================] - 0s 5ms/step - loss: 2.1344 - accuracy: 0.3824

Epoch 59/200

2/2 [==============================] - 0s 6ms/step - loss: 2.0989 - accuracy: 0.3824

Epoch 60/200

2/2 [==============================] - 0s 6ms/step - loss: 2.0641 - accuracy: 0.3824

Epoch 61/200

2/2 [==============================] - 0s 6ms/step - loss: 2.0409 - accuracy: 0.3824

Epoch 62/200

2/2 [==============================] - 0s 4ms/step - loss: 2.0243 - accuracy: 0.4118

Epoch 63/200

2/2 [==============================] - 0s 5ms/step - loss: 2.0030 - accuracy: 0.4118

Epoch 64/200

2/2 [==============================] - 0s 5ms/step - loss: 1.9711 - accuracy: 0.4118

Epoch 65/200

2/2 [==============================] - 0s 5ms/step - loss: 1.9392 - accuracy: 0.4118

Epoch 66/200

2/2 [==============================] - 0s 6ms/step - loss: 1.9070 - accuracy: 0.4412

Epoch 67/200

2/2 [==============================] - 0s 5ms/step - loss: 1.8823 - accuracy: 0.4412

Epoch 68/200

2/2 [==============================] - 0s 4ms/step - loss: 1.8635 - accuracy: 0.4118

Epoch 69/200

2/2 [==============================] - 0s 5ms/step - loss: 1.8362 - accuracy: 0.4118

Epoch 70/200

2/2 [==============================] - 0s 4ms/step - loss: 1.8038 - accuracy: 0.5000

Epoch 71/200

2/2 [==============================] - 0s 4ms/step - loss: 1.7620 - accuracy: 0.5000

Epoch 72/200

2/2 [==============================] - 0s 4ms/step - loss: 1.7261 - accuracy: 0.4412

Epoch 73/200

2/2 [==============================] - 0s 5ms/step - loss: 1.6860 - accuracy: 0.5000

Epoch 74/200

2/2 [==============================] - 0s 4ms/step - loss: 1.6592 - accuracy: 0.5000

Epoch 75/200

2/2 [==============================] - 0s 4ms/step - loss: 1.6409 - accuracy: 0.5294

Epoch 76/200

2/2 [==============================] - 0s 4ms/step - loss: 1.6168 - accuracy: 0.5000

Epoch 77/200

2/2 [==============================] - 0s 4ms/step - loss: 1.5851 - accuracy: 0.4706

Epoch 78/200

2/2 [==============================] - 0s 5ms/step - loss: 1.5540 - accuracy: 0.5294

Epoch 79/200

2/2 [==============================] - 0s 4ms/step - loss: 1.5164 - accuracy: 0.5294

Epoch 80/200

2/2 [==============================] - 0s 4ms/step - loss: 1.4713 - accuracy: 0.5588

Epoch 81/200

2/2 [==============================] - 0s 4ms/step - loss: 1.4381 - accuracy: 0.5588

Epoch 82/200

2/2 [==============================] - 0s 5ms/step - loss: 1.4111 - accuracy: 0.5588

Epoch 83/200

2/2 [==============================] - 0s 5ms/step - loss: 1.3882 - accuracy: 0.6471

Epoch 84/200

2/2 [==============================] - 0s 4ms/step - loss: 1.3698 - accuracy: 0.6471

Epoch 85/200

2/2 [==============================] - 0s 4ms/step - loss: 1.3491 - accuracy: 0.6471

Epoch 86/200

2/2 [==============================] - 0s 4ms/step - loss: 1.3175 - accuracy: 0.6765

Epoch 87/200

2/2 [==============================] - 0s 4ms/step - loss: 1.2793 - accuracy: 0.6471

Epoch 88/200

2/2 [==============================] - 0s 4ms/step - loss: 1.2423 - accuracy: 0.6176

Epoch 89/200

2/2 [==============================] - 0s 5ms/step - loss: 1.2128 - accuracy: 0.6765

Epoch 90/200

2/2 [==============================] - 0s 4ms/step - loss: 1.1877 - accuracy: 0.6471

Epoch 91/200

2/2 [==============================] - 0s 5ms/step - loss: 1.1603 - accuracy: 0.6765

Epoch 92/200

2/2 [==============================] - 0s 6ms/step - loss: 1.1398 - accuracy: 0.7353

Epoch 93/200

2/2 [==============================] - 0s 4ms/step - loss: 1.1215 - accuracy: 0.6471

Epoch 94/200

2/2 [==============================] - 0s 4ms/step - loss: 1.1028 - accuracy: 0.6471

Epoch 95/200

2/2 [==============================] - 0s 4ms/step - loss: 1.0773 - accuracy: 0.7059

Epoch 96/200

2/2 [==============================] - 0s 4ms/step - loss: 1.0495 - accuracy: 0.7353

Epoch 97/200

2/2 [==============================] - 0s 5ms/step - loss: 1.0164 - accuracy: 0.7647

Epoch 98/200

2/2 [==============================] - 0s 4ms/step - loss: 0.9861 - accuracy: 0.7647

Epoch 99/200

2/2 [==============================] - 0s 4ms/step - loss: 0.9608 - accuracy: 0.7353

Epoch 100/200

2/2 [==============================] - 0s 5ms/step - loss: 0.9370 - accuracy: 0.7353

Epoch 101/200

2/2 [==============================] - 0s 4ms/step - loss: 0.9170 - accuracy: 0.7647

Epoch 102/200

2/2 [==============================] - 0s 5ms/step - loss: 0.8931 - accuracy: 0.7353

Epoch 103/200

2/2 [==============================] - 0s 5ms/step - loss: 0.8726 - accuracy: 0.7941

Epoch 104/200

2/2 [==============================] - 0s 4ms/step - loss: 0.8485 - accuracy: 0.8235

Epoch 105/200

2/2 [==============================] - 0s 4ms/step - loss: 0.8299 - accuracy: 0.8235

Epoch 106/200

2/2 [==============================] - 0s 4ms/step - loss: 0.8118 - accuracy: 0.7941

Epoch 107/200

2/2 [==============================] - 0s 5ms/step - loss: 0.7897 - accuracy: 0.7647

Epoch 108/200

2/2 [==============================] - 0s 4ms/step - loss: 0.7610 - accuracy: 0.8529

Epoch 109/200

2/2 [==============================] - 0s 4ms/step - loss: 0.7359 - accuracy: 0.8824

Epoch 110/200

2/2 [==============================] - 0s 4ms/step - loss: 0.7151 - accuracy: 0.8824

Epoch 111/200

2/2 [==============================] - 0s 5ms/step - loss: 0.7002 - accuracy: 0.8824

Epoch 112/200

2/2 [==============================] - 0s 5ms/step - loss: 0.6869 - accuracy: 0.8529

Epoch 113/200

2/2 [==============================] - 0s 5ms/step - loss: 0.6746 - accuracy: 0.8529

Epoch 114/200

2/2 [==============================] - 0s 5ms/step - loss: 0.6585 - accuracy: 0.8824

Epoch 115/200

2/2 [==============================] - 0s 4ms/step - loss: 0.6416 - accuracy: 0.8824

Epoch 116/200

2/2 [==============================] - 0s 4ms/step - loss: 0.6259 - accuracy: 0.8529

Epoch 117/200

2/2 [==============================] - 0s 5ms/step - loss: 0.6151 - accuracy: 0.8235

Epoch 118/200

2/2 [==============================] - 0s 5ms/step - loss: 0.6096 - accuracy: 0.8529

Epoch 119/200

2/2 [==============================] - 0s 5ms/step - loss: 0.5996 - accuracy: 0.8529

Epoch 120/200

2/2 [==============================] - 0s 5ms/step - loss: 0.5900 - accuracy: 0.8529

Epoch 121/200

2/2 [==============================] - 0s 5ms/step - loss: 0.5838 - accuracy: 0.8529

Epoch 122/200

2/2 [==============================] - 0s 4ms/step - loss: 0.5671 - accuracy: 0.9118

Epoch 123/200

2/2 [==============================] - 0s 4ms/step - loss: 0.5477 - accuracy: 0.9118

Epoch 124/200

2/2 [==============================] - 0s 5ms/step - loss: 0.5283 - accuracy: 0.9118

Epoch 125/200

2/2 [==============================] - 0s 5ms/step - loss: 0.5089 - accuracy: 0.9412

Epoch 126/200

2/2 [==============================] - 0s 5ms/step - loss: 0.4923 - accuracy: 0.9412

Epoch 127/200

2/2 [==============================] - 0s 4ms/step - loss: 0.4824 - accuracy: 0.9412

Epoch 128/200

2/2 [==============================] - 0s 6ms/step - loss: 0.4719 - accuracy: 0.9118

Epoch 129/200

2/2 [==============================] - 0s 4ms/step - loss: 0.4604 - accuracy: 0.9118

Epoch 130/200

2/2 [==============================] - 0s 4ms/step - loss: 0.4495 - accuracy: 0.9118

Epoch 131/200

2/2 [==============================] - 0s 4ms/step - loss: 0.4395 - accuracy: 0.9118

Epoch 132/200

2/2 [==============================] - 0s 5ms/step - loss: 0.4290 - accuracy: 0.8824

Epoch 133/200

2/2 [==============================] - 0s 4ms/step - loss: 0.4181 - accuracy: 0.8824

Epoch 134/200

2/2 [==============================] - 0s 4ms/step - loss: 0.4088 - accuracy: 0.8824

Epoch 135/200

2/2 [==============================] - 0s 4ms/step - loss: 0.4003 - accuracy: 0.9412

Epoch 136/200

2/2 [==============================] - 0s 5ms/step - loss: 0.3955 - accuracy: 0.9412

Epoch 137/200

2/2 [==============================] - 0s 4ms/step - loss: 0.3897 - accuracy: 0.8529

Epoch 138/200

2/2 [==============================] - 0s 4ms/step - loss: 0.3877 - accuracy: 0.8824

Epoch 139/200

2/2 [==============================] - 0s 4ms/step - loss: 0.3865 - accuracy: 0.8529

Epoch 140/200

2/2 [==============================] - 0s 5ms/step - loss: 0.3853 - accuracy: 0.9118

Epoch 141/200

2/2 [==============================] - 0s 4ms/step - loss: 0.3803 - accuracy: 0.9118

Epoch 142/200

2/2 [==============================] - 0s 4ms/step - loss: 0.3717 - accuracy: 0.9412

Epoch 143/200

2/2 [==============================] - 0s 4ms/step - loss: 0.3635 - accuracy: 0.9118

Epoch 144/200

2/2 [==============================] - 0s 4ms/step - loss: 0.3580 - accuracy: 0.9118

Epoch 145/200

2/2 [==============================] - 0s 5ms/step - loss: 0.3514 - accuracy: 0.8824

Epoch 146/200

2/2 [==============================] - 0s 4ms/step - loss: 0.3428 - accuracy: 0.8824

Epoch 147/200

2/2 [==============================] - 0s 4ms/step - loss: 0.3314 - accuracy: 0.9118

Epoch 148/200

2/2 [==============================] - 0s 4ms/step - loss: 0.3259 - accuracy: 0.9118

Epoch 149/200

2/2 [==============================] - 0s 5ms/step - loss: 0.3163 - accuracy: 0.8824

Epoch 150/200

2/2 [==============================] - 0s 5ms/step - loss: 0.3100 - accuracy: 0.9118

Epoch 151/200

2/2 [==============================] - 0s 4ms/step - loss: 0.3049 - accuracy: 0.9118

Epoch 152/200

2/2 [==============================] - 0s 4ms/step - loss: 0.2994 - accuracy: 0.9118

Epoch 153/200

2/2 [==============================] - 0s 4ms/step - loss: 0.2920 - accuracy: 0.9118

Epoch 154/200

2/2 [==============================] - 0s 4ms/step - loss: 0.2775 - accuracy: 0.9118

Epoch 155/200

2/2 [==============================] - 0s 5ms/step - loss: 0.2645 - accuracy: 0.9118

Epoch 156/200

2/2 [==============================] - 0s 4ms/step - loss: 0.2571 - accuracy: 0.9412

Epoch 157/200

2/2 [==============================] - 0s 4ms/step - loss: 0.2510 - accuracy: 0.9412

Epoch 158/200

2/2 [==============================] - 0s 4ms/step - loss: 0.2466 - accuracy: 0.9706

Epoch 159/200

2/2 [==============================] - 0s 4ms/step - loss: 0.2408 - accuracy: 0.9706

Epoch 160/200

2/2 [==============================] - 0s 4ms/step - loss: 0.2345 - accuracy: 1.0000

Epoch 161/200

2/2 [==============================] - 0s 4ms/step - loss: 0.2274 - accuracy: 1.0000

Epoch 162/200

2/2 [==============================] - 0s 4ms/step - loss: 0.2215 - accuracy: 1.0000

Epoch 163/200

2/2 [==============================] - 0s 4ms/step - loss: 0.2152 - accuracy: 1.0000

Epoch 164/200

2/2 [==============================] - 0s 4ms/step - loss: 0.2091 - accuracy: 1.0000

Epoch 165/200

2/2 [==============================] - 0s 5ms/step - loss: 0.2034 - accuracy: 1.0000

Epoch 166/200

2/2 [==============================] - 0s 4ms/step - loss: 0.1991 - accuracy: 1.0000

Epoch 167/200

2/2 [==============================] - 0s 4ms/step - loss: 0.1926 - accuracy: 1.0000

Epoch 168/200

2/2 [==============================] - 0s 5ms/step - loss: 0.1884 - accuracy: 1.0000

Epoch 169/200

2/2 [==============================] - 0s 5ms/step - loss: 0.1856 - accuracy: 1.0000

Epoch 170/200

2/2 [==============================] - 0s 5ms/step - loss: 0.1845 - accuracy: 1.0000

Epoch 171/200

2/2 [==============================] - 0s 4ms/step - loss: 0.1873 - accuracy: 0.9706

Epoch 172/200

2/2 [==============================] - 0s 5ms/step - loss: 0.1934 - accuracy: 0.9706

Epoch 173/200

2/2 [==============================] - 0s 5ms/step - loss: 0.1990 - accuracy: 0.9118

Epoch 174/200

2/2 [==============================] - 0s 4ms/step - loss: 0.2012 - accuracy: 0.9118

Epoch 175/200

2/2 [==============================] - 0s 5ms/step - loss: 0.1981 - accuracy: 0.9118

Epoch 176/200

2/2 [==============================] - 0s 4ms/step - loss: 0.1921 - accuracy: 0.9118

Epoch 177/200

2/2 [==============================] - 0s 5ms/step - loss: 0.1842 - accuracy: 0.9412

Epoch 178/200

2/2 [==============================] - 0s 5ms/step - loss: 0.1758 - accuracy: 0.9706

Epoch 179/200

2/2 [==============================] - 0s 4ms/step - loss: 0.1674 - accuracy: 0.9706

Epoch 180/200

2/2 [==============================] - 0s 4ms/step - loss: 0.1599 - accuracy: 0.9706

Epoch 181/200

2/2 [==============================] - 0s 4ms/step - loss: 0.1582 - accuracy: 0.9706

Epoch 182/200

2/2 [==============================] - 0s 4ms/step - loss: 0.1619 - accuracy: 0.9706

Epoch 183/200

2/2 [==============================] - 0s 5ms/step - loss: 0.1653 - accuracy: 0.9706

Epoch 184/200

2/2 [==============================] - 0s 5ms/step - loss: 0.1593 - accuracy: 0.9706

Epoch 185/200

2/2 [==============================] - 0s 4ms/step - loss: 0.1546 - accuracy: 0.9706

Epoch 186/200

2/2 [==============================] - 0s 4ms/step - loss: 0.1595 - accuracy: 0.9412

Epoch 187/200

2/2 [==============================] - 0s 4ms/step - loss: 0.1682 - accuracy: 0.9412

Epoch 188/200

2/2 [==============================] - 0s 5ms/step - loss: 0.1730 - accuracy: 0.9412

Epoch 189/200

2/2 [==============================] - 0s 4ms/step - loss: 0.1718 - accuracy: 0.9412

Epoch 190/200

2/2 [==============================] - 0s 4ms/step - loss: 0.1660 - accuracy: 0.9412

Epoch 191/200

2/2 [==============================] - 0s 5ms/step - loss: 0.1569 - accuracy: 0.9412

Epoch 192/200

2/2 [==============================] - 0s 4ms/step - loss: 0.1465 - accuracy: 0.9412

Epoch 193/200

2/2 [==============================] - 0s 5ms/step - loss: 0.1364 - accuracy: 0.9706

Epoch 194/200

2/2 [==============================] - 0s 4ms/step - loss: 0.1276 - accuracy: 0.9706

Epoch 195/200

2/2 [==============================] - 0s 5ms/step - loss: 0.1220 - accuracy: 0.9706

Epoch 196/200

2/2 [==============================] - 0s 5ms/step - loss: 0.1170 - accuracy: 0.9706

Epoch 197/200

2/2 [==============================] - 0s 4ms/step - loss: 0.1156 - accuracy: 0.9706

Epoch 198/200

2/2 [==============================] - 0s 4ms/step - loss: 0.1093 - accuracy: 1.0000

Epoch 199/200

2/2 [==============================] - 0s 5ms/step - loss: 0.1056 - accuracy: 1.0000

Epoch 200/200

2/2 [==============================] - 0s 4ms/step - loss: 0.1119 - accuracy: 0.9706

Out[15]:

<keras.src.callbacks.History at 0x7f9a47db7690>

In [16]:

*# Get the word embeddings*

embeddings **=** model**.**get\_weights()[0]

*# Perform PCA to reduce the dimensionality of the embeddings*

pca **=** PCA(n\_components**=**2)

reduced\_embeddings **=** pca**.**fit\_transform(embeddings)

**d. Output**

In [17]:

*# Visualize the embeddings*

plt**.**figure(figsize**=**(7, 7))

**for** i, word **in** enumerate(tokenizer**.**word\_index**.**keys()):

x, y **=** reduced\_embeddings[i]

plt**.**scatter(x, y)

plt**.**annotate(word, xy**=**(x, y), xytext**=**(5, 2),

textcoords**=**'offset points',

ha**=**'right', va**=**'bottom')

plt**.**show()

A diagram of a diagram

Description automatically generated with medium confidence

In [18]:

*# test model*

test\_sentenses **=** [

"we are to study",

"create programs direct processes",

"spirits process study program",

"idea study people create"

]

In [19]:

**for** test\_sentense **in** test\_sentenses:

test\_words **=** test\_sentense**.**split(" ")

print("Words: ", test\_words)

x\_test **=** []

**for** i **in** test\_words:

x\_test**.**append(word\_to\_index\_map**.**get(i))

x\_test **=** np**.**array([x\_test])

print("Indexs: ", x\_test)

test\_predictions **=** model**.**predict(x\_test)

y\_pred **=** np**.**argmax(test\_predictions[0])

print("Predictons: ",test\_words, " => ", index\_to\_word\_map**.**get(y\_pred))

print("\n")

Words: ['we', 'are', 'to', 'study']

Indexs: [[ 4 5 6 12]]

1/1 [==============================] - 0s 58ms/step

Predictons: ['we', 'are', 'to', 'study'] => about

Words: ['create', 'programs', 'direct', 'processes']

Indexs: [[33 34 35 3]]

1/1 [==============================] - 0s 14ms/step

Predictons: ['create', 'programs', 'direct', 'processes'] => to

Words: ['spirits', 'process', 'study', 'program']

Indexs: [[39 8 12 31]]

1/1 [==============================] - 0s 13ms/step

Predictons: ['spirits', 'process', 'study', 'program'] => are

Words: ['idea', 'study', 'people', 'create']

Indexs: [[13 12 32 33]]

1/1 [==============================] - 0s 13ms/step

Predictons: ['idea', 'study', 'people', 'create'] => programs