

# Main file

```
# nltk.download('punkt')
'''
The 'punkt' tokenizer is used for tokenizing text into individual words or sentences.
'''

# nltk.download('wordnet')
'''
The 'wordnet' corpus is a lexical database that provides synonyms, antonyms, and other word relationships.
'''

lemmatizer = WordNetLemmatizer()

# Load intents from a JSON file
intents = json.loads(open('intents2.json').read())

words = [] # List to store individual words
classes = [] # List to store classes (intent tags)
documents = [] # List to store word patterns with their corresponding intent tags
ignoreLetters = ['?', '!', '.', ','] # List of characters to ignore

# Extract words, classes, and word patterns from intents
for intent in intents['intents']:
    for pattern in intent['patterns']:
        wordList = nltk.word_tokenize(pattern) # Tokenize the pattern into words
        words.extend(wordList) # Add words to the words list
        documents.append((wordList, intent['tag'])) # Add word pattern and intent tag to documents list
        if intent['tag'] not in classes:
            classes.append(intent['tag']) # Add unique intent tags to classes list

# Lemmatize words (reduce them to their base form) and remove ignored characters
words = [lemmatizer.lemmatize(word) for word in words if word not in ignoreLetters]

# Sort and remove duplicates from words and classes lists
words = sorted(set(words))

classes = sorted(set(classes))
```

```
intents['intents']: [
{
'tag': 'greeting',
'patterns': ['Hi', 'How are you', 'Is anyone there?', 'Hello', 'Good day', 'Whats
up', 'Good morning', 'Good evening', 'hello', 'hey', "what's up"],
'responses': [
['Hello! How can I assist you today?', 'Good to see you! How may I help you?',
'Hi there, how can I assist you?'], 'context_set': "
}pattern
```

```
, {'tag': 'goodbye', 'patterns': ['cya', 'See you later', 'Goodbye', 'I am Leaving',
'Have a Good day', 'bye', 'i have to go', 'gotta go'], 'responses': ['Sad to see you
go. Have a great day!', 'Talk to you later. Take care!', 'Goodbye! Have a
wonderful day!'], 'context_set': ''}, {'tag': 'age', 'patterns': ['how old', 'how old
are you', 'what is your age', 'how old are you', 'age?'], 'responses': ["I am a
virtual assistant, so I don't have an age!", "I'm an AI-powered bot, so age
doesn't apply to me!"], 'context_set': ''}, {'tag': 'name', 'patterns': ['what is your
name', 'what should I call you', 'whats your name?', 'who are you?'],
'responses': ['You can call me CallBot.', "I'm CallBot!", "I'm your friendly
CallBot."], 'context_set': ''}, {'tag': 'help', 'patterns': ['Id like to ask something',
'what can you do', 'can you help me?', 'can i tell you something'], 'responses':
["I'm here to help you! How can I assist you today?", 'I can help you with a wide
range of inquiries. What do you need assistance with?', "I'm here to assist you.
Please let me know how I can help."], 'context_set': ''}, {'tag':
'customer_service', 'patterns': ['I have a question about my order', 'I need
assistance with a product', 'Can you help me with a billing issue?', 'I want to
provide feedback', 'I need technical support', 'Can you transfer me to a live
agent?'], 'responses': ['Sure, I can assist you with that. Please provide me with
more details about your question or issue.', "Of course! I'm here to help. Please
let me know the specific problem or question you have.", "I'll do my best to
assist you. Please provide me with more information about your request."],
'context_set': ''}] ,
```

type: <class 'list'>

```
=====
=====
```

wordList: ['Can', 'you', 'transfer', 'me', 'to', 'a', 'live', 'agent', '?'] ,

type: <class 'list'>

```
=====
=====
```

classes: ['greeting', 'goodbye', 'age', 'name', 'help', 'customer\_service'] ,

type: <class 'list'>

=====

words: ['Hi', 'How', 'are', 'you', 'Is', 'anyone', 'there', '?', 'Hello', 'Good', 'day', 'Whats', 'up', 'Good', 'morning', 'Good', 'evening', 'hello', 'hey', 'what', "'s", 'up', 'cya', 'See', 'you', 'later', 'Goodbye', 'I', 'am', 'Leaving', 'Have', 'a', 'Good', 'day', 'bye', 'i', 'have', 'to', 'go', 'got', 'ta', 'go', 'how', 'old', 'how', 'old', 'are', 'you', 'what', 'is', 'your', 'age', 'how', 'old', 'are', 'you', 'age', '?', 'what', 'is', 'your', 'name', 'what', 'should', 'I', 'call', 'you', 'whats', 'your', 'name', '?', 'who', 'are', 'you', '?', 'Id', 'like', 'to', 'ask', 'something', 'what', 'can', 'you', 'do', 'can', 'you', 'help', 'me', '?', 'can', 'i', 'tell', 'you', 'something', 'I', 'have', 'a', 'question', 'about', 'my', 'order', 'I', 'need', 'assistance', 'with', 'a', 'product', 'Can', 'you', 'help', 'me', 'with', 'a', 'billing', 'issue', '?', 'I', 'want', 'to', 'provide', 'feedback', 'I', 'need', 'technical', 'support', 'Can', 'you', 'transfer', 'me', 'to', 'a', 'live', 'agent', '?'] , type: <class 'list'>

=====

documents: [

(['Hi'], 'greeting'), (['How', 'are', 'you'], 'greeting'), (['Is', 'anyone', 'there', '?'], 'greeting'), (['Hello'], 'greeting'), (['Good', 'day'], 'greeting'), (['Whats', 'up'], 'greeting'), (['Good', 'morning'], 'greeting'), (['Good', 'evening'], 'greeting'), (['hello'], 'greeting'), (['hey'], 'greeting'), (['what', "'s", 'up'], 'greeting'), (['cya', 'goodbye'], 'goodbye'), (['See', 'you', 'later'], 'goodbye'), (['Goodbye'], 'goodbye'), (['I', 'am', 'Leaving'], 'goodbye'), (['Have', 'a', 'Good', 'day'], 'goodbye'), (['bye', 'goodbye'], 'goodbye'), (['i', 'have', 'to', 'go'], 'goodbye'), (['got', 'ta', 'go'], 'goodbye'), (['how', 'old'], 'age'), (['how', 'old', 'are', 'you'], 'age'), (['what', 'is', 'your', 'age'], 'age'), (['how', 'old', 'are', 'you'], 'age'), (['age', '?'], 'age'), (['what', 'is', 'your', 'name'], 'name'), (['what', 'should', 'I', 'call', 'you'], 'name'), (['whats', 'your', 'name', '?'], 'name'), (['who', 'are', 'you', '?'], 'name'), (['Id', 'like', 'to', 'ask', 'something'], 'help'), (['what', 'can', 'you', 'do'], 'help'), (['can', 'you', 'help', 'me',

```
'?'], 'help'), (['can', 'i', 'tell', 'you', 'something'], 'help'), (['I', 'have', 'a', 'question',
'about', 'my', 'order'], 'customer_service'), (['I', 'need', 'assistance', 'with', 'a',
'product'], 'customer_service'), (['Can', 'you', 'help', 'me', 'with', 'a', 'billing',
'issue', '?'], 'customer_service'), (['I', 'want', 'to', 'provide', 'feedback'],
'customer_service'), (['I', 'need', 'technical', 'support'], 'customer_service'),
(['Can', 'you', 'transfer', 'me', 'to', 'a', 'live', 'agent', '?'], 'customer_service')
],
```

type: <class 'list'>

```
classes.index(document[1])
classes.index(['Hi']) will return the index
```

```
=====

words after Lemmatize: ['Hi', 'How', 'are', 'you', 'Is', 'anyone', 'there', 'Hello',
'Good', 'day', 'Whats', 'up', 'Good', 'morning', 'Good', 'evening', 'hello', 'hey',
'what', "'s", 'up', 'cya', 'See', 'you', 'later', 'Goodbye', 'I', 'am', 'Leaving', 'Have',
'a', 'Good', 'day', 'bye', 'i', 'have', 'to', 'go', 'got', 'ta', 'go', 'how', 'old', 'how',
'old', 'are', 'you', 'what', 'is', 'your', 'age', 'how', 'old', 'are', 'you', 'age', 'what',
'is', 'your', 'name', 'what', 'should', 'I', 'call', 'you', 'whats', 'your', 'name', 'who',
'are', 'you', 'Id', 'like', 'to', 'ask', 'something', 'what', 'can', 'you', 'do', 'can', 'you',
'help', 'me', 'can', 'i', 'tell', 'you', 'something', 'I', 'have', 'a', 'question', 'about',
'my', 'order', 'I', 'need', 'assistance', 'with', 'a', 'product', 'Can', 'you', 'help',
'me', 'with', 'a', 'billing', 'issue', 'I', 'want', 'to', 'provide', 'feedback', 'I', 'need',
'technical', 'support', 'Can', 'you', 'transfer', 'me', 'to', 'a', 'live', 'agent'] type:
```

<class 'list'>

```
=====

words set after Lemmatize : ["'s", 'Can', 'Good', 'Goodbye', 'Have', 'Hello', 'Hi',
'How', 'I', 'Id', 'Is', 'Leaving', 'See', 'Whats', 'a', 'about', 'age', 'agent', 'am',
'anyone', 'are', 'ask', 'assistance', 'billing', 'bye', 'call', 'can', 'cya', 'day', 'do',
'evening', 'feedback', 'go', 'got', 'have', 'hello', 'help', 'hey', 'how', 'i', 'is', 'issue',
'later', 'like', 'live', 'me', 'morning', 'my', 'name', 'need', 'old', 'order', 'product',
'provide', 'question', 'should', 'something', 'support', 'ta', 'technical', 'tell',
```

'there', 'to', 'transfer', 'up', 'want', 'what', 'whats', 'who', 'with', 'you', 'your']

type: <class 'list'>

=====

Classes list after Lemmatize : ['greeting', 'goodbye', 'age', 'name', 'help',  
'customer\_service']

type: <class 'list'>

=====

classes set after Lemmatize : ['age', 'customer\_service', 'goodbye', 'greeting',  
'help', 'name']

type: <class 'list'>

=====

```
# Create training data by converting word patterns to bag of words format and generating output rows
# preparing the data for training a model.
for document in documents:
    bag = []
    wordPatterns = document[0]

    wordPatterns = [lemmatizer.lemmatize(word.lower()) for word in wordPatterns]
    for word in words:
        bag.append(1) if word in wordPatterns else bag.append(0)

    outputRow = list(outputEmpty)

    outputRow[classes.index(document[1])] = 1

    training.append(bag + outputRow)

# Shuffle the training data randomly
random.shuffle(training)
"""
By shuffling the training data, you are changing the order of the elements randomly.
This can be useful to ensure that the data is not biased by any specific order during training.
Shuffling the data helps to introduce randomness and prevent the model from learning any patterns that may exist due to the order of the data.
"""
```

outputRow: [0, 0, 0, 0, 0, 0] type: <class 'list'>

outputRow a: [0, 1, 0, 0, 0, 0] type: <class 'list'>

```
wordPatterns: ['Can', 'you', 'help', 'me', 'with', 'a', 'billing', 'issue', '?'] type:
<class 'list'>
```

=====

[illegible]

[illegible]





[illegible]

[illegible]



train\_y: [

[0 1 0 0 0 0] [0 0 0 0 1 0] [0 0 1 0 0 0] [0 0 1 0 0 0] [1 0 0 0 0 0] [0 1 0 0 0 0] [0 0 1  
0 0 0] [0 0 1 0 0 0] [1 0 0 0 0 0] [0 1 0 0 0 0] [0 0 0 0 1 0] [0 0 1 0 0 0] [0 0 0 0 1 0]  
[0 0 0 0 0 1] [0 0 0 1 0 0] [1 0 0 0 0 0] [0 0 0 1 0 0] [0 0 0 1 0 0] [0 0 0 1 0 0] [0 0  
0 0 0 1] [0 1 0 0 0 0] [0 0 0 1 0 0] [1 0 0 0 0 0] [0 0 0 1 0 0] [0 0 0 1 0 0] [0 0 0 0 1  
0] [0 0 0 1 0 0] [0 1 0 0 0 0] [1 0 0 0 0 0] [0 0 1 0 0 0] [0 0 0 1 0 0] [0 0 0 0 0 1] [0  
0 0 1 0 0] [0 1 0 0 0 0] [0 0 1 0 0 0] [0 0 0 1 0 0] [0 0 0 0 0 1] [0 0 1 0 0 0]],

type: <class 'numpy.ndarray'>

```
=====
model = tf.keras.Sequential()
# Add layers to the model
model.add(tf.keras.layers.Dense(128, input_shape=(len(trainX[0]),), activation='relu')) # Input layer
"""
: This line adds the input layer to the model. The Dense layer represents a fully connected layer,
where each neuron is connected to every neuron in the previous layer. The 128 parameter specifies the number of neurons in this layer.
The input_shape parameter defines the shape of the input data, which in this case is (len(trainX[0]),).
The activation='relu' parameter specifies the activation function to be used, which is the Rectified Linear Unit (ReLU) in this case.
"""
model.add(tf.keras.layers.Dropout(0.5)) # Dropout layer for regularization
"""
This line adds a dropout layer to the model.
Dropout is a regularization technique that randomly sets a fraction of input units to 0 during training,
which helps prevent overfitting.
The 0.5 parameter specifies the dropout rate, which is the fraction of input units to drop.
"""
model.add(tf.keras.layers.Dense(64, activation='relu')) # Hidden layer
"""
This line adds a hidden layer to the model. It is similar to the input layer, but with 64 neurons instead of 128.
"""
model.add(tf.keras.layers.Dropout(0.5)) # Dropout layer for regularization
"""
This line adds another dropout layer for regularization, similar to the previous one.
"""
model.add(tf.keras.layers.Dense(len(trainY[0]), activation='softmax')) # Output layer
"""
This line adds the output layer to the model. The number of neurons in this layer is determined by the number of classes in your target variable,
which is len(trainY[0]) in this case. The activation='softmax' parameter specifies the activation function to be used, which is the softmax function.
Softmax is commonly used for multi-class classification problems as it produces a probability distribution over the classes.
"""
```

2023-10-30 13:00:17.042285: 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: SSE SSE2 SSE3 SSE4.1 SSE4.2 AVX AVX2 FMA, in other operations, rebuild TensorFlow with the appropriate compiler flags.

---

```
sgd = tf.keras.optimizers.SGD(learning_rate=0.01, momentum=0.9, nesterov=True)

"""
This line creates an instance of the Stochastic Gradient Descent (SGD) optimizer.
The SGD optimizer is a popular optimization algorithm used in neural networks.
The learning_rate parameter determines the step size at each iteration of the optimization process.
The momentum parameter controls the amount of momentum to apply during optimization, which helps the optimizer to converge faster.
The nesterov parameter specifies whether to use Nesterov momentum, which is an improvement over standard momentum.
"""

model.compile(loss='categorical_crossentropy', optimizer=sgd, metrics=['accuracy'])

"""
This line compiles the model. The loss parameter specifies the loss function to be used during training.
In this case, 'categorical_crossentropy' is used, which is commonly used for multi-class classification problems.
The optimizer parameter specifies the optimizer to be used for training, which is the sgd optimizer that we created earlier.
The metrics parameter specifies the evaluation metric(s) to be used during training and testing.
In this case, we are using 'accuracy' as the metric to monitor the performance of the model.
"""

# Train the model using the training data
hist = model.fit(np.array(trainX), np.array(trainY), epochs=200, batch_size=5, verbose=1)

"""
epochs=200: This parameter specifies the number of times the entire training dataset will be passed through the model during training.
In this case, the model will be trained for 200 epochs.

batch_size=5: This parameter determines the number of samples that will be propagated through the model at once.
The training dataset will be divided into batches,
and each batch will be used to update the model's weights. A smaller batch size can help with memory constraints and can also lead to faster convergence,
but it may introduce more noise in the training process.

verbose=1: This parameter controls the verbosity of the training process. Setting it to 1 will display progress bars during training,
while setting it to 0 will suppress the output
The model.fit function will train the model by iteratively updating the model's weights based on the provided input and target data.
It uses the specified optimizer, loss function, and evaluation metric to guide the training process.
After training is complete, the fit function will return a History object, which contains information about the training process,
such as the loss and accuracy values at each epoch.
If you have any further questions or need more clarification, feel free to ask!
"""
```

Example of fitting:

Batch\_size: how much of words

Epochs: no of iterations of training

This will return because show\_metrics = true

Epoch 1/200

8/8 [=====] - 1s 1ms/step - loss: 1.8203 - accuracy: 0.1579

Epoch 2/200

8/8 [=====] - 0s 2ms/step - loss: 1.7833 - accuracy: 0.1842

Epoch 3/200

Type of the model:

type:<class 'keras.src.callbacks.History'>

```
# Save words and classes lists using pickle
pickle.dump(words, open('words.pkl', 'wb'))
pickle.dump(classes, open('classes.pkl', 'wb'))

"""
This line saves the words variable to a file named 'words.pkl'.
The pickle.dump() function serializes the object (words in this case) and writes it to the file.
The 'wb' argument specifies that the file should be opened in binary mode for writing.
"""
```

Done