**ChatBot**

**Bag-of-Words (BoW)**

* Popular technique in natural language processing (NLP)
* Machine learning for representing text data
* Convert textual information into a numerical format that can be used by machine learning algorithms

**Step 1: Tokenization**

* The first step is to break down a piece of text into individual words or tokens.
* Removing punctuation
* Converting all characters to lowercase
* Split the text into words

**Step 2: Stemming**

* Is an NLP technique
* Information retrieval to reduce words to their base or root form

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

**Pattern**

A diagram of a network

Description automatically generated

**Softmax**

* Softmax is a mathematical function
* Used as the activation function in the output layer to convert the raw output scores into probabilities.

**ReLU (Rectified Linear Unit)**

* An activation function commonly used in neural networks and deep learning models
* Allowing it to learn complex patterns and relationships in the data

**To start developing Chatbot**

1. Create Environment for Python

python3 -m venv venv

1. Activate Environment

venv\Scripts\activate

1. Libraries need to be installed:

* pip install nltk (Natural Language Toolkit)
* ntlk.download(‘punkt’)
* pip install numpy (Numerical Python)
* pip install torch \*Currently only support Python 3.8 - 3.11\*
* pip install jieba (Chinese text segmentation)
* pip install langid (Language Identification)

1. Stored messages or data in Json file

**References**

<https://www.youtube.com/watch?v=RpWeNzfSUHw&ab_channel=PatrickLoeber>

<https://github.com/patrickloeber/pytorch-chatbot/tree/master>