West Visayas State University

COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Luna St., La Paz, Iloilo City 5000
Iloilo, Philippines
* Trunkline: (063) (033) 320-0870 loc 1403   * Telefax No.: (033) 320-0879
* Website: www.wvsu.edu.ph   * Email Address: cict@wvsu.edu.ph

Exercise for Unit 2

Name: LEANN VILLARUEL & DONJIE LIBUNA

Date: 02/21/2026

Year and Section: BSCS 3B - AI

Note: Write your answers in a one whole sheet of paper.

1. **Updating ELIZA.** Modify the ELIZA python implementation shared to the class by performing the following:
   (10 points each) Add 5 RegEx patterns to capture :
   a. I want to know the reasons why I am feeling depressed all the time.
   b. I am feeling stressed.
   c. My feelings towards my crush are invalidated.
   d. You don't understand me OR You do not understand me.
   e. I can't focus on my studies OR I cannot focus on my studies.

   BONUS! (10 points) Modify the code so that if the user tries to repeat the same question, ELIZA will answer sarcastically on repeating the question.

2. **Updating ELIZA. Implementing RegEx on NLP.** Perform the following preprocessing techniques using RegEx.

   a. (10 points) Extract all of the *words starting with an upper case letter* from the text:

   Alice was beginning to get very tired of sitting by her sister on the bank, and of having nothing to do. Once or twice she had peeped into the book her sister was reading, but it had no pictures or conversations in it, "and what is the use of a book," thought Alice, "without pictures or conversations?

West Visayas State University

COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Luna St., La Paz, Iloilo City 5000
Iloilo, Philippines
* Trunkline: (063) (033) 320-0870 loc 1403   * Telefax No.: (033) 320-0879
* Website: www.wvsu.edu.ph   * Email Address: cict@wvsu.edu.ph

RegEx pattern: r'\b[A-Z][a-zA-Z]*\b'

b. (10 points) Read the "melville-moby_dick.txt" into a Python program and extract all of the instances of the word *whale* in said source.

RegEx pattern: r'\bwhales?\b'

c. (10 points) Download the NLTK package into your local environment and import *webtext (from nltk.corpus import webtext).* Load the pirates.txt and extract all of the lines spoken by the character Jack Sparrow.

RegEx pattern: r'^JACK SPARROW:\s*(.*)$'