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 7

Name: **Mauricio Manuel F. Bergancia** Date: **May 16, 2025**

Name: **Mherlie Joy U. Chavez**

Year and Section: **BSCS 3-A AI**

Note: Upload your sources in a public GitHub repository and put the link in the document.

1. (70 points) Choose two NLP tasks:
   a. **PoS Tagging**
   b. NER Recognition
   c. Text Prediction
   d. **Sentiment Classification**

   Create 2 models, 1) the traditional (machine learning technique appropriate for the task) used, and 2) train a neural network (Feedforward, RNN, LSTM, also appropriate for the task).

   For the corpus, you can use any appropriate dataset for the NLP Task. You can use Python libraries suited for neural networks.

- In this exercise, we chose the following topics with the following models:

a. **Sentiment Classification** - using the IMDb dataset, which contains 50,000 movie reviews labeled as **positive** or **negative**.
   - **Naive Bayes** - A traditional machine learning classifier trained on TF-IDF features. It assumes word independence and works well for simple text tasks.
   - **LSTM Neural Network** - A deep learning model using an embedding layer followed by an LSTM, trained on tokenized and padded sequences. It can capture word order and context over time.

   Both models were trained on the same training data and evaluated on the same test set of 25,000 reviews.

b. **PoS Tagging** - also known as grammatical tagging, is a fundamental process in Natural Language Processing.
   - **Hidden Markov Model** - a statistical model that estimates the probability of a sequence of tags given a sequence of words.
   - **LSTM Neural Network** - trained on padded sequences of tokenized words and their corresponding POS tags, using one-hot encoding for tags.

2. (30 points) Evaluate each model and compare its performance. Which of the models has performed the best?

a. Sentiment Classification

Metrics used:

Metric	How it is used
Accuracy	The percentage of total predictions the model got right.
Precision	Of the predicted positives, how many were actually positive? (Low precision = more false positives)
Recall	Of all actual positives, how many did the model correctly predict? (Low recall = more false negatives)
F1-score	A balance between precision and recall. Good for uneven class distributions or when both false positives and negatives matter.
Support	Number of actual occurrences of each class in the test set.

Results:

Metric	Naive Bayes	LSTM Neural Net	Better Model
Accuracy	0.83	**0.85**	LSTM
Positive F1	0.84	**0.86**	LSTM
Negative F1	0.82	**0.85**	LSTM
Macro Avg F1	0.83	**0.85**	LSTM

The **LSTM neural network** is the better-performing model for this sentiment analysis task based on accuracy, precision, recall, and F1-score. While Naive Bayes is fast and simple, LSTM captures richer patterns in the data — which is especially useful in understanding sentiment from text.

b. PoS Tagging

Metric used:

Accuracy - The percentage of POS tags predicted **correctly** over the total number of tags. In both models, accuracy is computed by comparing predicted vs. actual tags for each word across the entire test dataset.

Results:

Model	Accuracy
HMM	**91.50%**
LSTM	**93.43%**

The **LSTM model** outperformed the HMM model by nearly 2 percentage points in accuracy. **LSTM** is the better model when it comes to **accuracy and performance** for POS tagging.

3. Put the GitHub repo link here: https://github.com/mherlie/CCS249-EXERCISE-UNIT-7