This is individual work. Collaborations are not allowed.

Please use Python to complete this homework. The requirements.txt includes most of what you need. First create a virtual environment. For example,

python3 -m venv .env

would create a (hidden) directory called .env under the current directory. Then, update pip using

python3 -m pip install --upgrade pip

Then install the packages from requirements.txt as in

pip install -r requirements.txt

To install additional packages do:

pip install \*package\_name\*

Then start jupyterlab as:

jupyter lab

Then, open the file: unigramLG.ipynb

This file shows how to generate sentences of fixed length (10 words) using a unigram model.

The corpus for your homework is in: 5 papers related to blockchains.docx

Process this data as appropriate (for example, creating a txt file) and read the file in a Jupyter notebook. Then, use the functions from nltk to develop language generators for bigram, trigram and 4-gram. Using each generator to generate 5 sentences of 10 words each.

Develop your code as a Jupyter notebook. Run your code and upload the code after running without resetting the Jupyter notebook kernel.

Start early and ask questions as you encounter difficulties.

**References.**

<https://anaconda.org/conda-forge/docx2txt>