Python IV – Lesson 17

Date: Oct 2, 2022

Agenda

- ► Google Colab
- ► Pandas

Proverbs 4:27

- "Do not turn to the right or the left; Remove your foot from evil. "

  - Abhor what is evil; hold fast to what is good.
  - Whoever guards his way preserves his life.
  - An evil man is ensnared by the transgression of his lips.

Homework: data_structure.py

1. We have a list of tuples, e.g. [(0, 1), (0, 9), (5, 7), (2, 4)]
   - ► write a function to find the tuple with the largest sum
   - ► Each tuple in the list should have 2 numbers in it. The function should return the tuple with the biggest second number.

1. We have a letters_to_numbers dictionary:
   - ► write a function that will take a string as an argument and use the letters_to_numbers dictionary to replace any letters it can with numbers
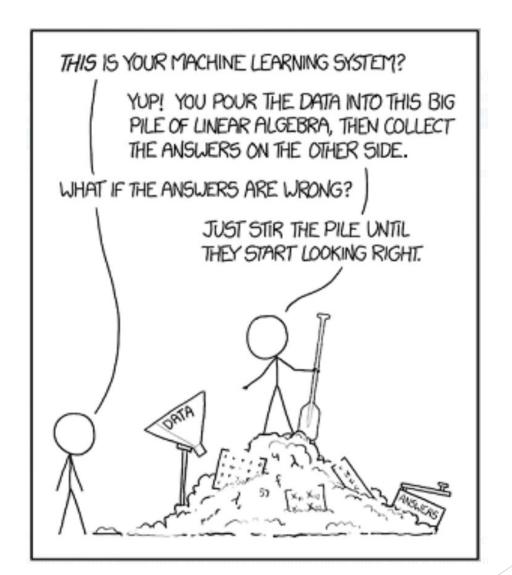
```
letters_to_numbers = {
    'a': 0,
    'b': 1,
    'c': 2,
    'd': 3,
    'e': 4
}
```

1. Write a function that returns the longest word from a string. (A word is any sequence of characters separated by whitespace)

Pandas

A library that you can't avoid when working with Python on a data science project.

Pandas contains a large variety of functions for data import, export, indexing, and data manipulation.

Terms you need to know

DataFrames — series of columns and rows

Series — 1-dimensional arrays and efficient methods for handling them.

Data Wrangling — steps required to prepare the data so that it can actually be consumed for extracting insights and model building.

Continue from lesson16:

1. Learn from
   https://colab.research.google.com/notebooks/mlcc/intro_to_pandas.ipynb#scrollTo=aSRYu62xUi3g
2. Work on
   https://colab.research.google.com/drive/1GwzDroONxUIMq4sKkoMRTeZqNDwH1yMV#scrollTo=g4gjBobo74sb

Let's learn Pandas by creating a new .ipynb file

- Series
  - From Lists
  - From Dictionaries
- DataFrame
  - From a Series Object
  - From a Dictionary
  - From a File
- Data Frame Operations

Data Source:

https://www.kaggle.com/datasets/PromptCloudHQ/imdb-data