BDS 761 – Practice Exam I

Name: _____

Time: 120 minutes

- If I cannot find or understand your writing then I cannot give you credit.
- If you do not show your work you may not get full credit.
- If you get the correct answer but add incorrect information you may not get full credit.
- *If your answer is too vague, or not the best answer, you also may not get full credit.*
- If a question asks for pseudocode, this means a simplified programming language you can invent yourself with an obvious syntax
- **CLOSED BOOK. CLOSED NOTES. NO WRITTEN SUPPORT MATERIAL ALLOWED. NO COMPUTERS. NO CELL/SMART PHONES.**

1. (5 points) BONUS.

   - Write your full name on this page plus at the top of every scratch/loose page you use (if you remove the staple your name must be on every separate page; it is ok to remove the staple).
   - Do not write on the backs of any pages.
   - Do not fold any pages.

2. (5 points) What are the two alternative types of string matching and how do they differ? Both functionally as well as in terms of pros and cons.

3. (15 points) Describe how you could perform a comparison between two documents that is lower (better) when they contain similar words, and higher when they don't. Give steps with as much detail as possible using code or pseudocode. We would like singular, plural, and other versions of a word to be treated as the same word.

4. (5 points) Explain the difference between dynamic programming and memoization. Which approach requires you to rewrite the algorithm? Suppose you have a slow function that can be accelerated by these methods. In what situations would you *not* want to use the accelerated version, even though it could provide a speed benefit?

5. (3 points) Give the mathematical equation for a dot product for vectors with *N* elements each

   (3 points) How many computations will this require?

   (10 points) Suppose we know the vectors each have at most *k* elements which are nonzero. We further know the indices of these nonzero elements along with their values. Describe an efficient program for computing the dot product much faster than the above, by using this information. How many computations will it take?

6. (3 points) Give the mathematical equation for matrix-vector multiplication

7. (5 points) Give the mathematical equation vector-matrix multiplication, i.e., for multiplying a matrix by a vector from the left side instead of the right side.

8. (4 points) Give the general mathematical equation for matrix-matrix multiplication

9. (10 points) Describe a way to store a m by n matrix in python. Give code for accessing a specific matrix element A_{ij} from this data structure.

10. (10 points) Give code for a function that takes a single input N and returns a N by N identity matrix.

11. (5 points) Give two reasons why a python set is not a good choice for describing vectors. What built-in data structure is best?

12. (10 points) Given two vectors with N elements each, we could use a dot product or a Euclidean distance to compare their similarity. Explain mathematically why these two approaches are closely related. What will be the differences between these two comparison values?

13. (15 points) List (and describe each mathematically) all the possibly ways you can think of to "multiply" two vectors with N elements each (including with broadcasting) using numpy. Including different order and possible transposes.