CS 712 - Special Topics in AI
Course Project Specification Weightage: 27%
Submission Deadline: 30-April-2023

General Instructions:

- You must use only C, C++, Python or JAVA for this course project.
- Prescribed specifications must be strictly followed. Failure to do so may lead to substantial loss of points.
- Make sure your code is well written (self explanatory variable names) and documented. You are likely to lose points if your TA cannot understand your code.

Question 1 (100 points): For this course project, you are required to develop and implement a variation of the spatial co-location pattern mining algorithm which works on the modified cross-K function given below. Following is a formal definition of the problem defining its input, output and objective.

Modified Cross-K Function:

$$CrossKF(f1, f2,..., fk) = \frac{(|f1\,Join\,f2\,Join\,f2\,....\,Join\,fk|)\ under\ the\ distance\ threshold\ h}{(|f1| \times |f2| \times |f3| \times ....|fk|)}$$

Input: (a) a set **F** of k spatial boolean feature types; **(b)** event instances for each feature type in **F**. Each instance is associated with a latitude and longitude; **(c)** neighborhood distance threshold **h;** **(d)** minimum area threshold **MinA**; **(e)** maximum area threshold **MaxA;** **(f)** interest measure threshold theta.

Output: A set of patterns P. Following holds for each pattern p_i ∈ **P**

- ☐ Interest measure of p_i is greater than theta.
- ☐ Boundary information of region Rp_i where p_i is prevalent. Area of Rp_i is greater than **MinA** and less than **MaxA**.
- ☐ p_i is maximal in Rp_i in the sense that any pattern which is a superset of p_i does not cross the threshold of theta.