CAPITOLO 4

Spiego il database, i vari attributi. La frequenza dei crimini per quanto riguarda il database preso su kaggle (o del 2018?) basarsi su quello visto su kaggle!!! FATTO

Spiego le scelte di taglio effettuate al dataset FATTO

Spiego i dataset utilizzati DA FARE (faccio nel quinto?)

scaletta delle robe da dire: divisione delle classi, sptree come funziona a grandi linee e dettaglio su struttura nodo con riguardo al dizionario, codice stbfm cerca di rispecchiare l'algoritmo ma alcune cose sono custom come la distanza spaziale tramite gps, ricorda che non sono possibili tipi ripetuti all'intero di una sequenza e le difficoltà trovate

input – output da riga di comando

Racconto brevemente il codice, del tipo le classi e i principali metodi, input – output da tastiera, schematizzare bene come ho fatto l’albero spiegando il perché e il fatto che ho utilizzato i dizionari

CAPITOLO 5

Analizzo i risultati sul dataset2018 e cerco di trarne conclusioni, possibilmente utilizzando grafici

Più il dataset aumenta meglio è scegliere un raggio spaziale e temporale minori, in quanto se il raggio e grande prenderà molti più valori di una certa tipologia rispetto ad un dataset di minore grandezza. Se invece il raggio è minore si hanno risposte più sensate e con tempi di computazione minore

Conclusioni e prospettive future