

Машинное обучение.

Домашнее задание №2

Задача 1. Покажите, что расстояние Джаккарда является метрикой.

Подсказка: может пригодиться доказанное на занятии утверждение о том, что коэффициент Джаккарда равен кое-какой вероятности.

Задача 2. Пусть зависимость вероятности коллизии от расстояния между объектами имеет вид $p = 1 - d$ (для исходного семейства хэш-функций). После построения композиции над этим семейством вероятность примет вид $1 - (1 - (1 - d)^m)^L$. Данное выражение как функция от p имеет вид сигмоиды. Найдите координату точки ее перегиба, и исследуйте ее поведение при стремлении L к бесконечности при фиксированном m (и наоборот, при стремлении m к бесконечности при фиксированном L). Проинтерпретируйте результаты.

Задача 3. Рассмотрим двумерное пространство ($d = 2$). Покажите, что описанное на занятии семейство хэш-функций для евклидовой метрики является $(r/2, 2r, 1/2, 1/3)$ -чувствительным.