

1. В чем разница между полнотой и точностью?

Точность(*precision*) показывает, какая доля объектов, выделенных классификатором как положительные, действительно является положительными

Полнота(*recall*) показывает, какая часть положительных объектов была выделена классификатором.

TP – True positive

FP – False positive

FN – False negative

$$precision = \frac{TP}{TP + FP}$$

$$recall = \frac{TP}{TP + FN}$$

2. Какие основные 3 фазы есть в ART сетях?

Инициализация сети – устанавливается параметр сходства $R_{кр}$ ($0 < R_{кр} < 1$), причем чем он больше, тем выше должно быть сходство образа и кластера для отнесения к последнему. Для первого поданного на входы сети образа создается первый нейрон (кластер)

Распознавание образа – на входы сети подается новый входной образ, для каждого нейрона рассчитывается значение его выхода с учетом коэффициентов кратковременной памяти. Положительные выходы нейронов указывают на кластеры, имеющие качественное сходство с входным образом

Кластеризация образа (сравнение) – рассчитывается количественная мера сходства входного образа с кластером, имеющим

наибольшее значение выхода нейрона. Если входного образа с кластером, имеющим наибольшее значение выхода нейрона больше параметра сходства, то входной образ соответствует этому кластеру. Если условие не выполняется, алгоритм повторяется. Если нейрон-победитель не найден, создается новый кластер с синаптическими коэффициентами нейрона.

3. Какую метрику можно использовать для сравнения двух текстов?

Можно использовать расстояние Левенштейна, которое представляет собой минимальное количество операций вставки, удаления и замены одного символа на другой, необходимых для того, чтобы превратить одну строку в другую.