

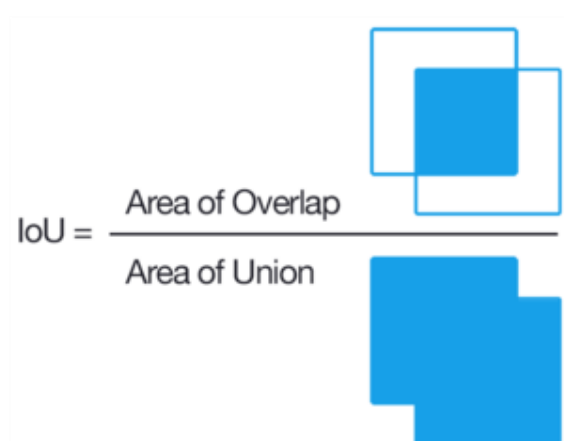
## 1. Что такое ART-сеть?

Сеть адаптивного резонанса для распознавания, хранения и классификации образов, работающая как в форме двоичных сигналов (ART-1), так и в форме аналоговых (ART-2). Для последнего типа возможно и комбинирование обеих форм сигналов в одной архитектуре. «Жизненный цикл» таких сетей не имеет строгого деления на стадии обучения, тестирования и эксплуатации. Это все связано с тем, что ART-сети самоорганизующиеся, т. е. на момент создания модель не имеет выходных нейронов – они создаются сами по мере надобности в ходе эксплуатации (сеть способна сама выяснить то, по скольким классам необходимо выполнить ранжирование входного сигнала). ART-сети имеют 2 типа синапсов (нейронных связей): долговременной и кратковременной памяти. Кратковременная память определяет кластеры, к которым может принадлежать входной образ, долговременная же выявляет степени соответствия образа к каждому из них (в некоторых случаях принимает решение о создании нового класса).

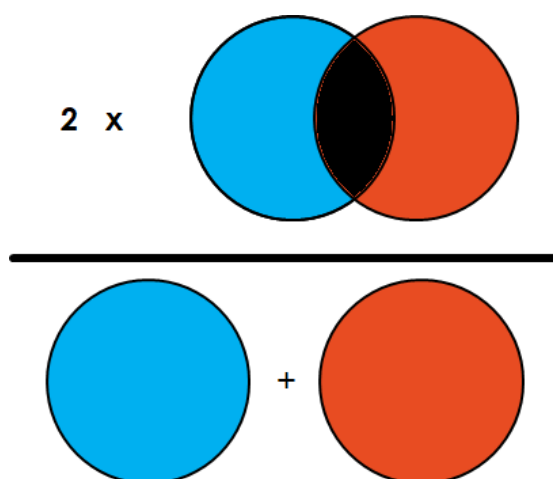
## 2. Какую метрику можно использовать для сравнения двух текстов?

Для примера можно привести два подхода:

IoU – Intersection-Over-Union, или Jaccard Index. Метод основан на нахождении отношения количества совпавших слов к общему кол-ву слов в двух текстах.



Второй способ похож на этот – Dice Coefficient. Отличие в умножении числителя на 2, что приближает результат к 1 (а оба коэфф-та измеряются от 0 до 1).



Оба подхода достаточно просты и распространены, хотя и имеют ярко выраженный недостаток в сильной зависимости результата от размеров сравниваемых текстов. Точность вычислений будет выше, чем больше слов окажется в тестах. Также стоит отметить, что при сравнении и работе с текстами стоит немного «подготовить» выборки, приняв разные склонения и некоторые формы слов (зависящие от подлежащего, например, в английском языке окончания) за одинаковые. Также лучше будет исключать из сравниваемых выборок т. н. стоп-слова – предлоги, союзы и пр. лексемы, не играющие особой роли в смысловом наполнении текстов.

### 3. Почему обучение длилось 2 эпохи?

В ходе исследования было выявлено, что 2 эпохи обучения – самое оптимальное решение в том плане, что при большем количестве наблюдается переобучение модели. Также обучение начинает занимать значительно больше времени.

