

1. Какую проблему решают ART-сети?

Для многих видов сетей при добавлении образца к обучающей выборке требуется заново обучать сеть на новой обучающей выборке. Это чаще всего долгая операция, потому что, во-первых, обучающая выборка чаще всего большая, а во-вторых, архитектура сетей часто бывает сложной, и обучение вычислительно дорого. Это называется проблемой стабильности/пластичности ('plasticity/stability' problem). ART-сети решают ее следующим образом: сеть 'помнит' набор образцов, для которых ей известен верный ответ, и сравнивает их с образцом, для которого требуется сделать предсказание. В результате сравнения вычисляется мера отличия, которая сравнивается с параметром бдительности ('vigilance parameter'). Если мера отличия меньше этого параметра, то образец классифицируется как принадлежащий соответствующей категории. Чтобы добавить образец в обучающую выборку, не требуется переобучать сеть, так как ей надо просто запомнить образцы (набор его признаков) и его категорию.

2. Что такое слой Гроссберга?

В сетях встречного распространения есть слои Кохонена и слои Гроссберга. Слой Кохонена работает так, что лишь у одного нейрона выход равен 1, а у остальных 0. Вот как работает слой Гроссберга: на вход подается вектор $K = (k_1, \dots, k_n)$, а выход $Y_j = \sum_i k_i w_{ij}$, где w_{ij} – элемент матрицы весов ребер, соединяющих нейроны Кохонена с нейронами Гроссберга. Поскольку лишь один нейрон Кохонена дает ненулевой единичный выход, каждый нейрон Гроссберга выдает значение веса ребра, соединяющего этот нейрон Гроссберга с нейроном Кохонена, имеющим единичный выход.

3. Почему в качестве функций активации на скрытых слоях используется ReLU?

Известно, что использование этой функции активации почти всегда дает неплохой результат. Кроме того, она вычислительно проще, чем гиперболический тангенс или сигмоида, поэтому сеть обучится быстрее, а нам ведь нужно обучать ее несколько раз.