

## 1. Что такое слой Гроссберга?

Слой Гроссберга — слой сети встречного распространения, который обучается следующим образом: его входной вектор является выходом слоя Кохонена (каждый нейрон в слое Кохонена соединен с каждым нейроном в слое Гроссберга). На этот вектор подается на слой нейронов Гроссберга, и выход является взвешенной суммой выходов слоя Кохонена. Причем, каждый вес корректируется только если он соединен с нейроном Кохонена, имевшим ненулевой выход.

$$Y_i = \sum_j k_j w_{ij},$$

где  $Y_i$  — выход  $j$ -го нейрона Гроссберга,

или в векторной форме  $Y = KV$ ,

где  $K$  — выходной вектор слоя Кохонена,

$V$  — матрица весов слоя Гроссберга.

Фактически каждый нейрон слоя Гроссберга выдает величину веса, который связывает этот нейрон с нейроном слоя Кохонена, являющимся единственным с величиной выхода, равной единице (у слоя Кохонена только у одного нейрона величина выхода равна единице, а у остальных — нулю).

## 2. В каких задачах применяется метрика F-мера?

F measure (F1 score, F score) является показателем точности теста. Определяется как взвешенное среднее гармоническое от precision и recall теста:

$$2 \cdot \frac{precision \cdot recall}{precision + recall}$$

Очень наглядно объясняют, что такое precision и recall следующие формулы:

$$precision = \frac{true\ positives}{true\ positives + false\ positives} = \frac{terrorists\ correctly\ identified}{terrorists\ correctly\ identified + individuals\ incorrectly\ labeled\ as\ terrorists}$$

$$recall = \frac{true\ positives}{true\ positives + false\ negatives} = \frac{terrorists\ correctly\ identified}{terrorists\ correctly\ identified + terrorists\ incorrectly\ labeled\ as\ not\ terrorists}$$

То есть recall выражает способность находить все релевантные экземпляры в наборе данных, precision выражает пропорцию данных, которые наша модель назвала релевантными, и они действительно были релевантными.

Обычно F-мера используется для оценки качества: очистки изображений и улучшения изображений древних документов. Эта задача состоит в оценке вероятности расположения чернил в каждом пикселе очищаемого изображения учитывая шумный аналог. Использование F-меры в купе с системой методов улучшения изображений древних документов применяется не только для улучшения читаемости таких документов для людей, но также является первой стадией в большинстве процессов предварительной обработки, применяемых в распознавании текста.

### 3. В чем особенность функции LReLU?

Функция ReLU задаётся формулой:  $f(x) = \begin{cases} x, & \text{if } x > 0 \\ 0 & \end{cases}$ .

При использовании такой активационной функции не происходит вымывания градиента, т.к. частная производная этой функции равна 1, когда  $x > 0$ . Однако при оптимизации появляются проблемы, заключающиеся в слишком медленном обучении в случаях с константным нулевым градиентом или в случае с вымыванием градиента. Из-за того, что градиент равен нулю в том числе, когда нейрон неактивен, результатом может стать то, что нейрон никогда не будет активирован, и соответственно, градиентный алгоритм не произведет подстройки весов такого нейрона.

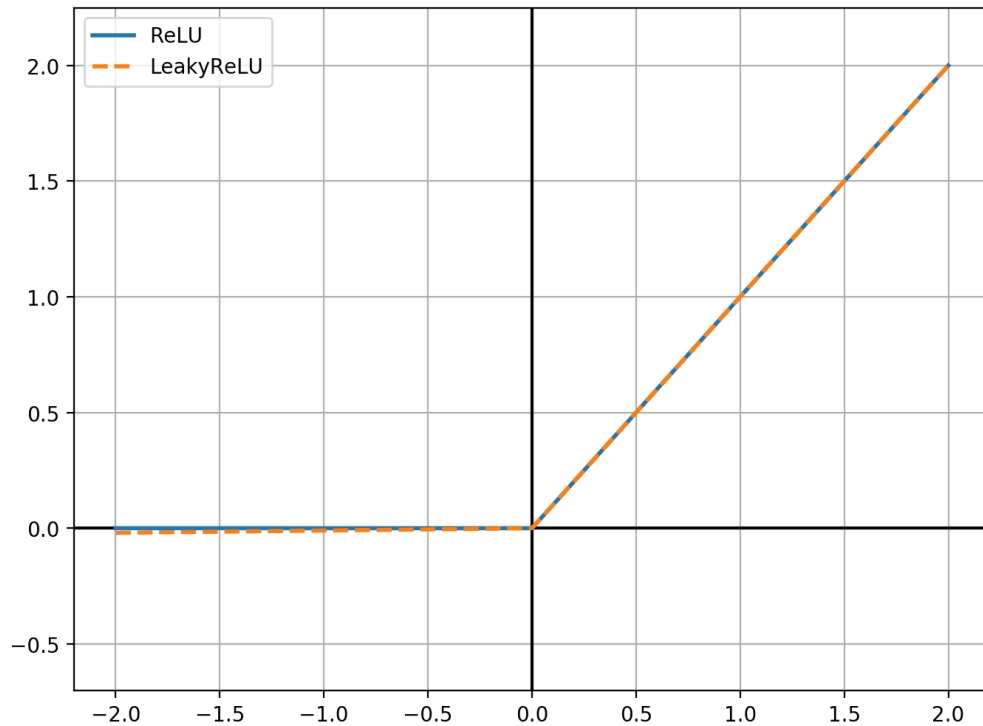
Для борьбы с этой проблемой используют модифицированную ReLU: Leaky ReLU

$$f(x) = \begin{cases} x, & \text{if } x > 0 \\ 0,01x & \end{cases}$$

. У нее будет маленький ненулевой градиент, когда

нейрон достигнет насыщения и станет неактивен.

На графике они очень похожи, но при  $x < 0$ , у LeakyReLU все же ненулевое значение. В этом ее особенность.



4. Почему в качестве функции активации на выходном слое используется сигмоид?

Основная причина, по которой мы используем сигмовидную функцию, заключается в том, что она существует между 0 и 1. Поэтому она особенно используется для моделей, в которых мы должны предсказать вероятность. Поскольку вероятность чего-либо существует только в диапазоне от 0 до 1, сигмоид является правильным выбором.

