



ОНЛАЙН-ОБРАЗОВАНИЕ

# Выучивание метрики.

Иногда расстояние мешает подобрать нужные слова

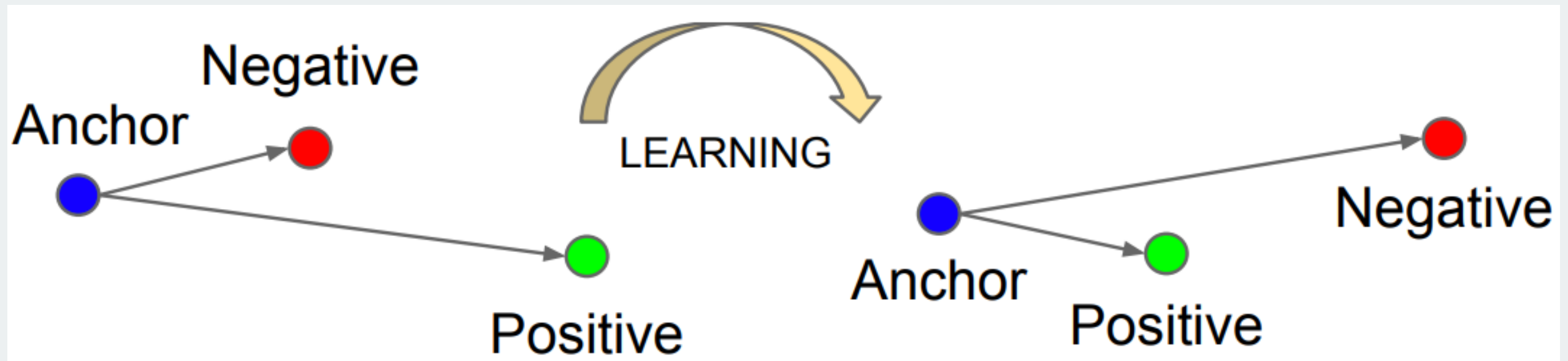
Артур Кадурын



1. **Опять тройка**
2. One-shot learning
3. Одно представление на всех



# Triplet loss

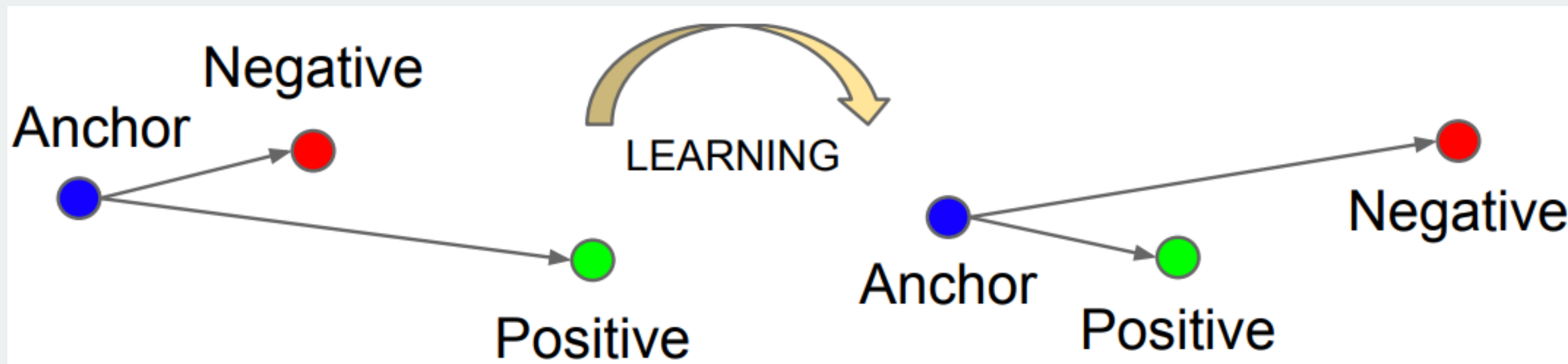


Предположим, что мы способны померять «расстояние» между объектами. Тогда, нам бы хотелось, чтобы объекты с одинаковыми или похожими метками были на небольшом расстоянии друг от друга и наоборот.

**Как должно быть устроено такое расстояние?**



# Triplet loss

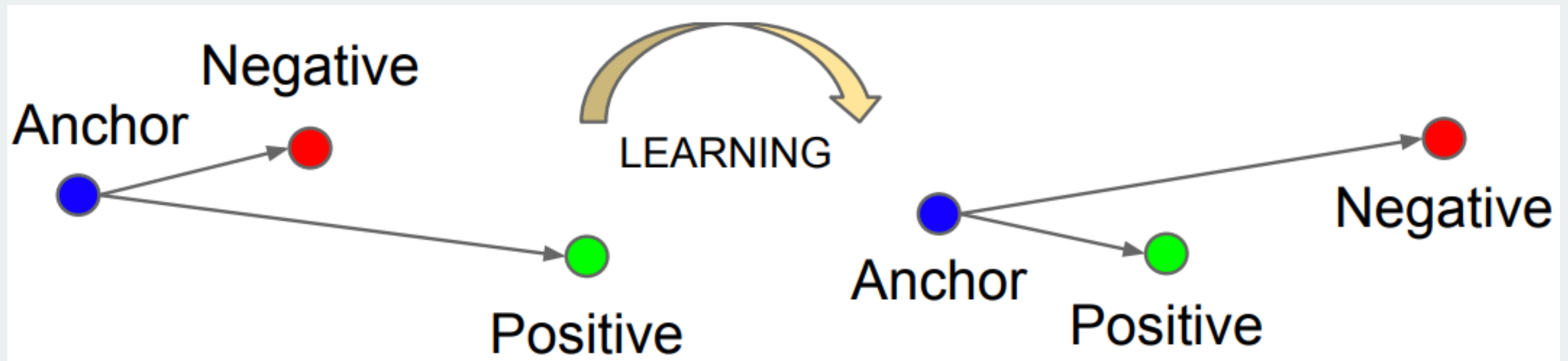


Допустим у нас есть нейросеть отображающая наши примеры в пространство над которым мы умеем мерять расстояние.

**Каким должен быть лосс?**



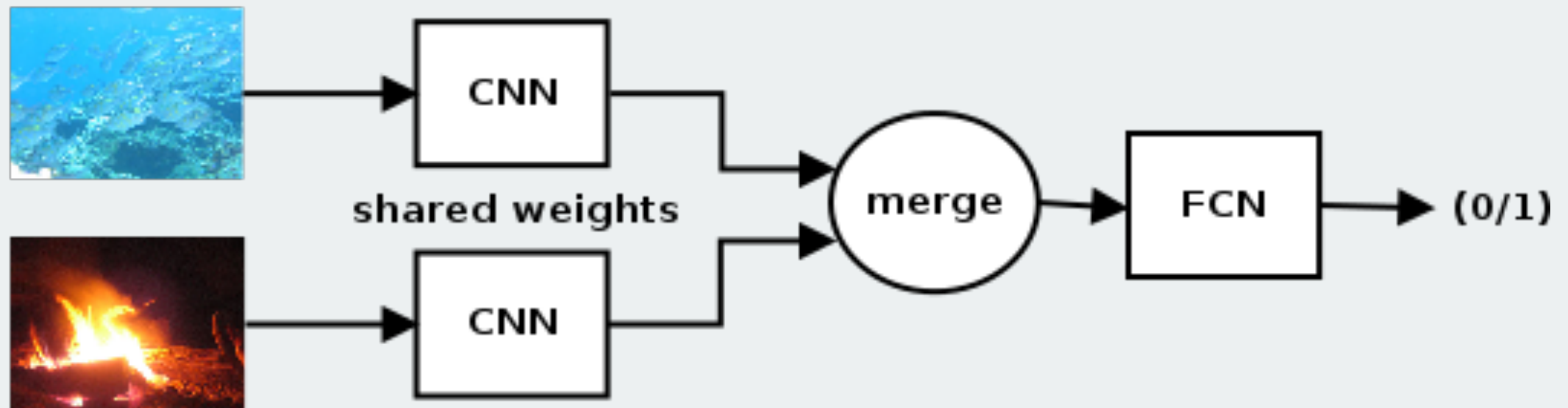
# Triplet loss



Допустим у нас есть нейросеть  $f: x \rightarrow R^d$  отображающая наши примеры в пространство над которым мы умеем мерять расстояние  $D$ .

$$L = \sum_i^N \left[ D \left( f(x_i^t), f(x_i^p) \right) - D \left( f(x_i^t), f(x_i^n) \right) + \alpha \right]_+$$



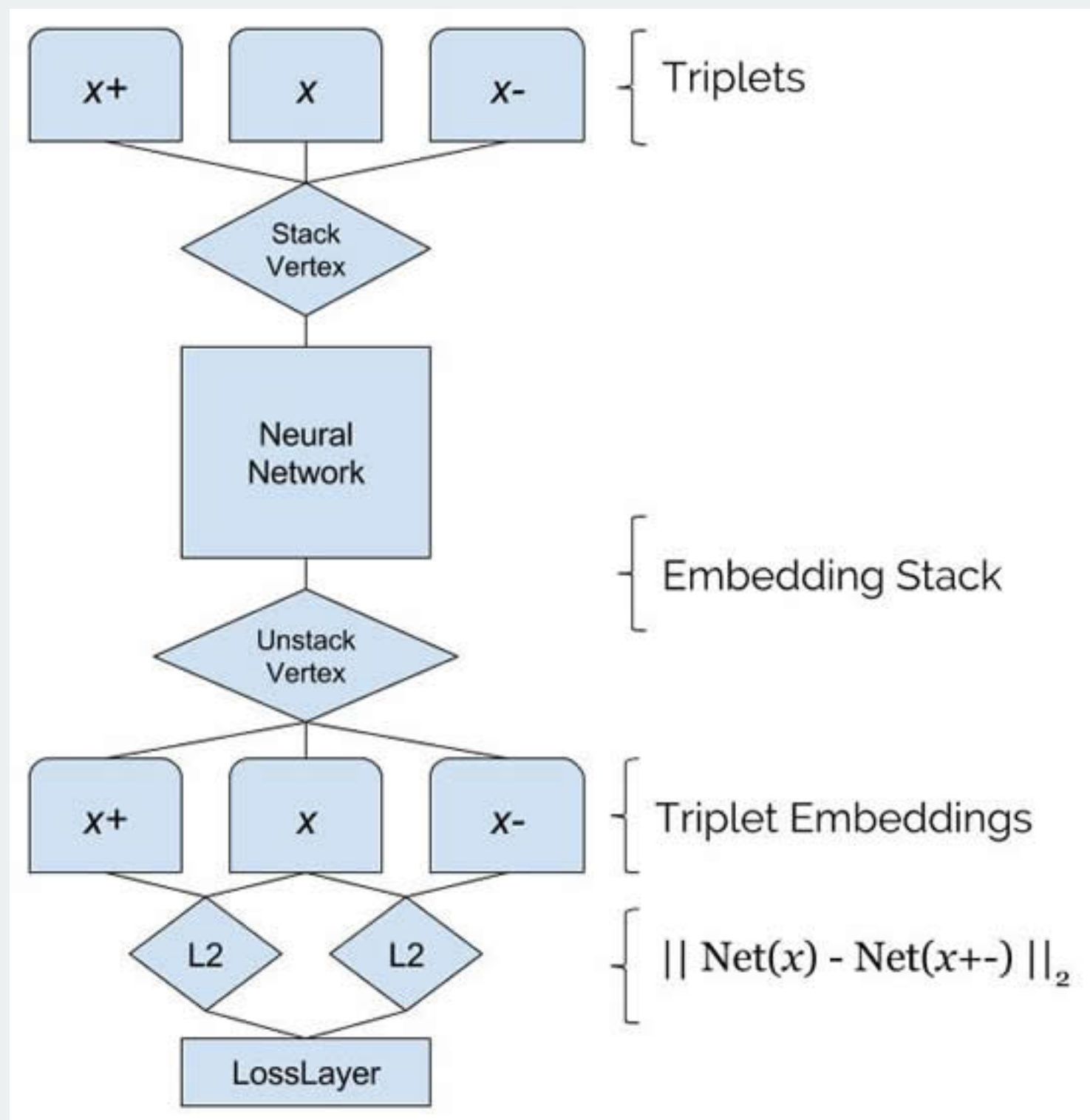


Сеть, которая сначала «сжимает» несколько примеров, а потом вычисляет над ними какую-нибудь функцию обычно называют сиамской.



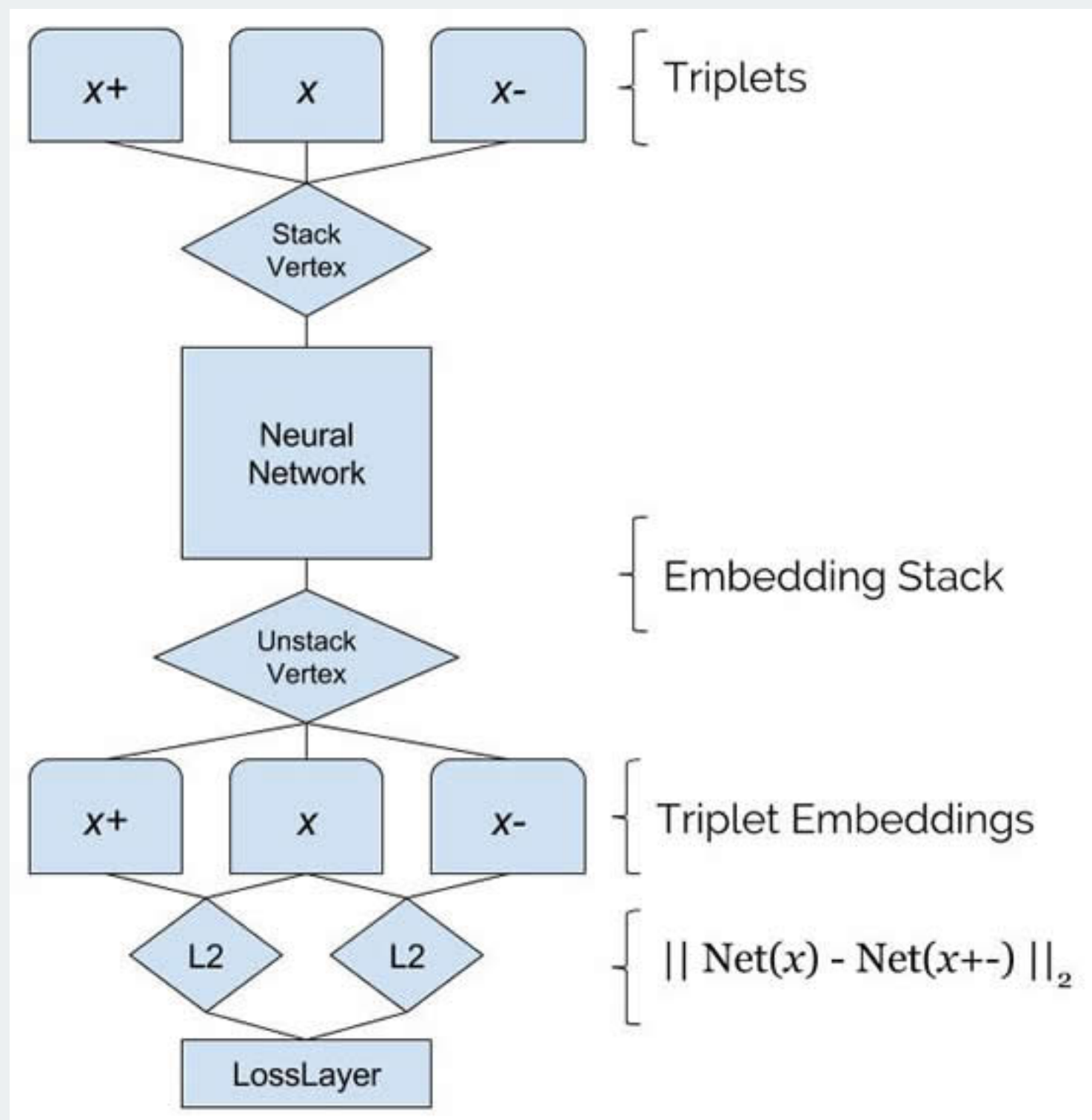
Триплеты собираются из троек: якорь, положительный пример, отрицательный пример.

При этом, есть различные стратегии выбора положительных и отрицательных примеров.





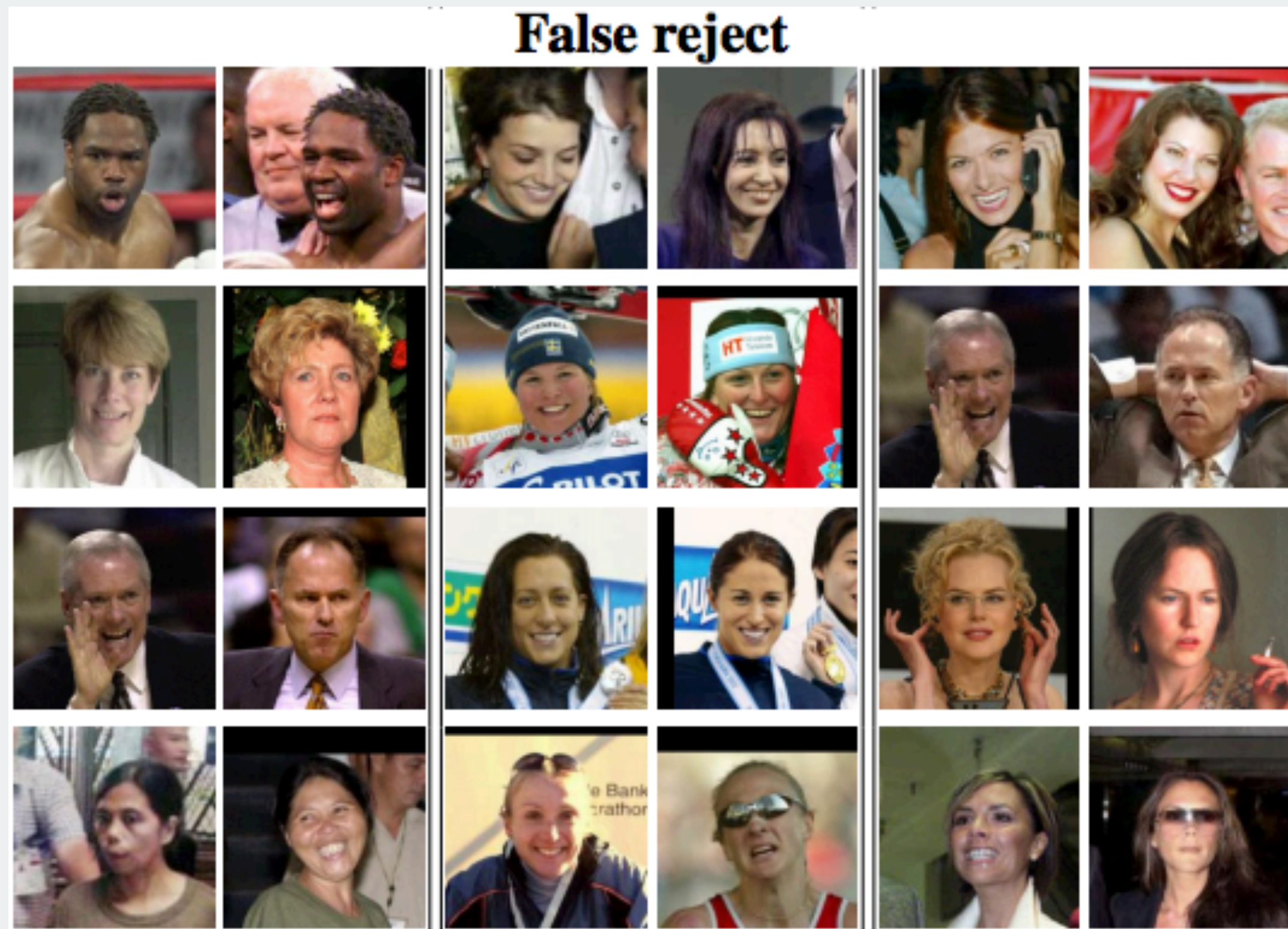
Кроме того, triplet loss может быть регуляризатором латентного слоя, а так же, вычисляться не только в последнем слое, но и после каждого слоя сети.



## False accept



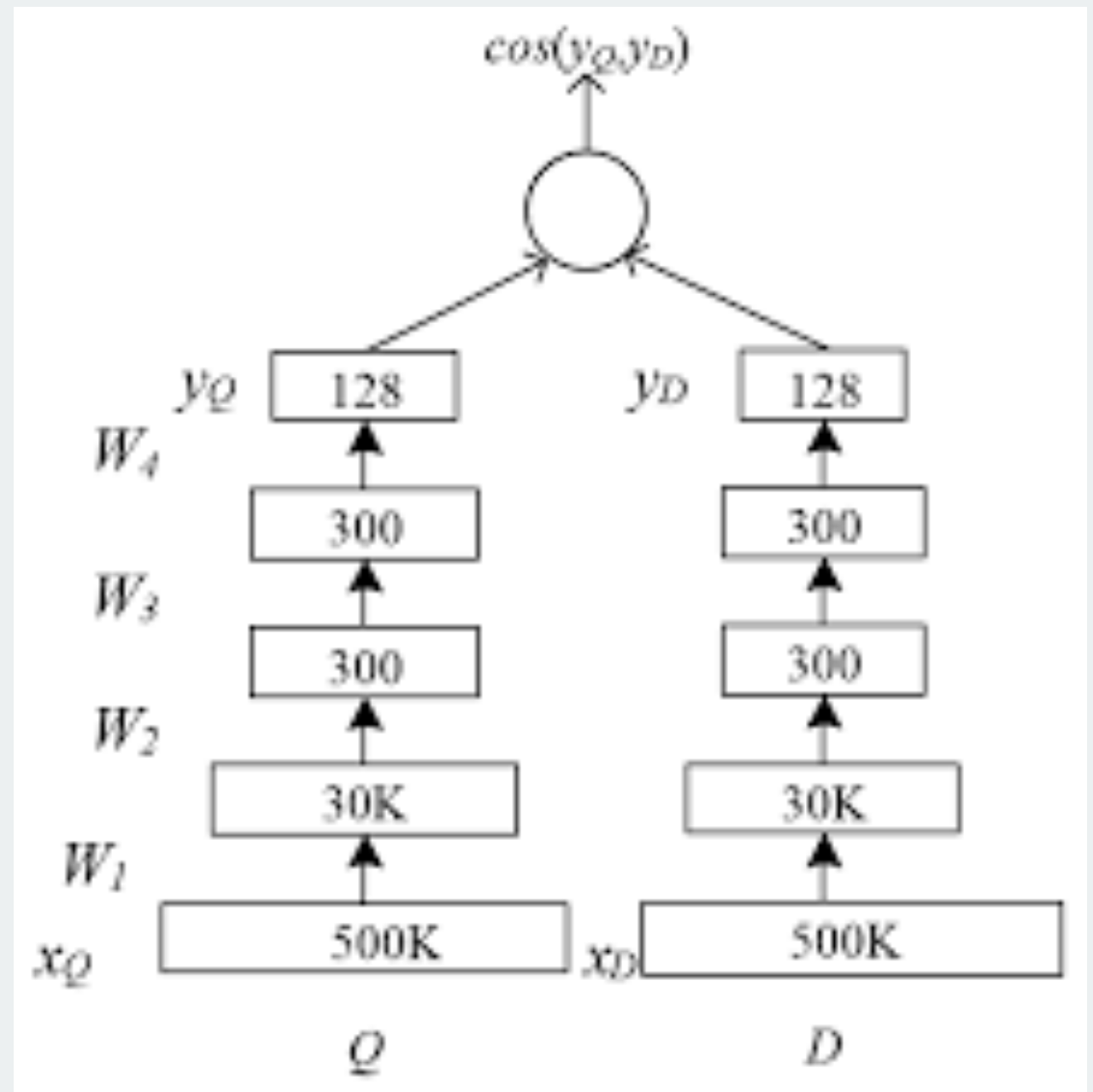




# Deep structured semantic model



В DSSM в качестве якоря используется запрос, а положительными и отрицательными примерами служат документы. При этом обе головы все так же используют одинаковые веса.



1. Опять тройка
2. **One-shot learning**
3. Одно представление на всех



[illegible]



[illegible]

Предсказание одного или нескольких тегов можно представить задачей классификации, однако, когда тегов сотни тысяч, собрать трейнсет для классификации практически невозможно.

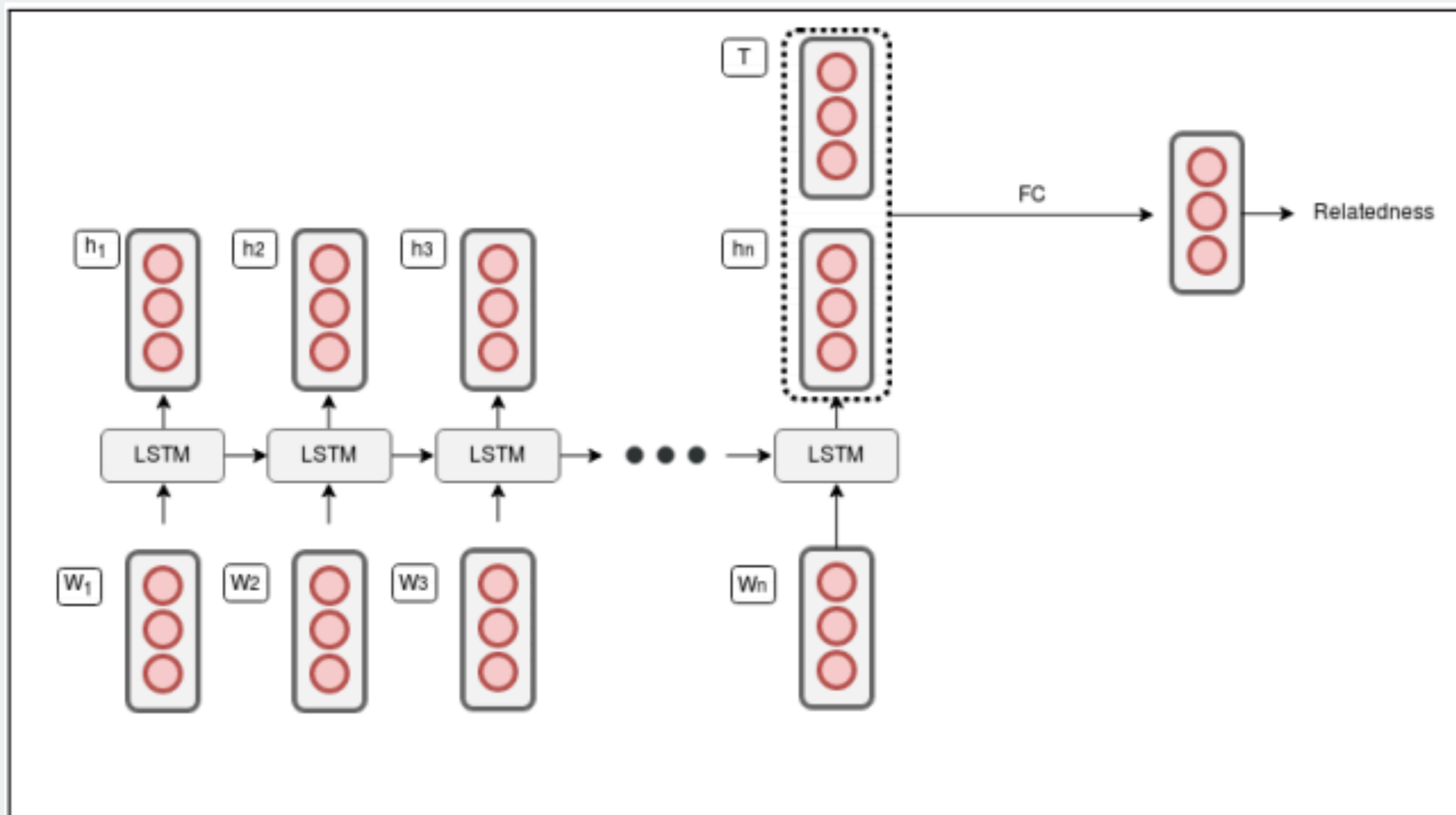
В статье от Paralleldots 2018 года задача классификации переформулирована в стиле выучивания метрик.

В качестве эмбедингов слов авторы использовали Google News

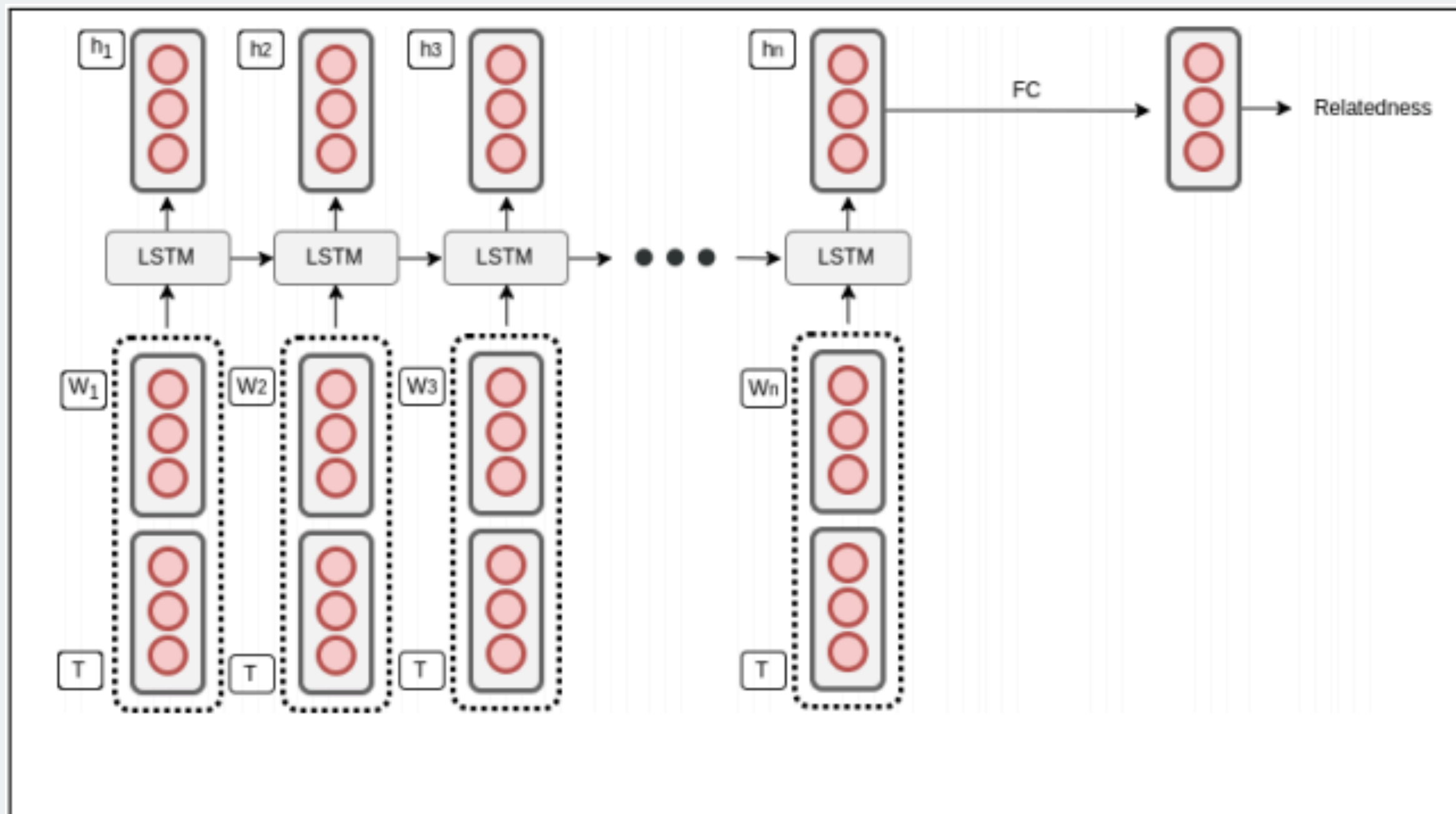
Embedding отсюда: <https://code.google.com/archive/p/word2vec/>





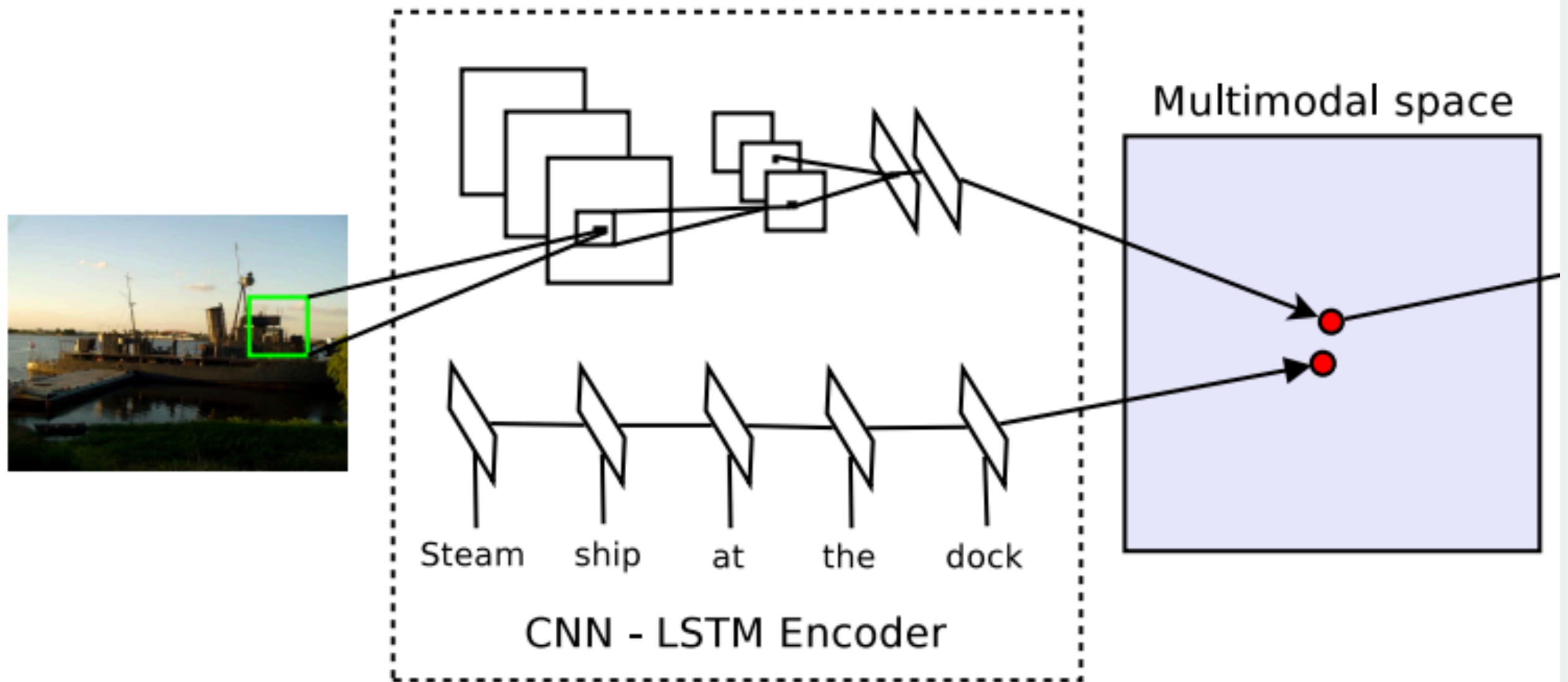


# Zero-shot тегирование



1. Опять тройка
2. One-shot learning
3. **Одно представление на всех**







Спасибо  
за внимание!