

Лабораторная работа 3

Студент: Будникова Валерия Павловна
Группа: М80-407Б-19

Статья 1: <https://trends.rbc.ru/trends/industry/64246fb99a7947ff3a55d05c>

В данной статье описывается, как работает сеть DALL-E. DALL-E - нейронная сеть, способная генерировать изображения по текстовому описанию. Идея этого подхода заключается в совмещении работы трех сетей - «CLIP (Contrastive Language–Image Pre-training), GLIDE и нейросети для увеличения разрешения картинки».

Данный подход можно рассмотреть с точки зрения non-parametric meta-learning. В начале работы сети, после первичной обработки текста, происходит процесс по выявлению связей и важности слов. Для этого применяется механизм внимания, который по характеристикам каждого слова во входном тексте определяет коэффициент суммы внимания. Далее, используя матрицу эмбедингов слов словаря, определяются коэффициенты каждого слова в словаре, в соответствии со словами из входной последовательности. На эту процедуру можно смотреть как на метод, делающий прогнозы, используя метрику сходства. Также, можно рассмотреть отдельно сеть CLIP. Во время обучения, при подаче на вход изображения, а на выход - формулировки предложений, в разной степени описывающих исходное изображение, предобученная CLIP на выходе получает численные метрики схожести каждой формулировки с изображением. Эта процедура отражает в себе подход классификации изображений.

Статья 2: <https://habr.com/ru/articles/599703/>

В статье описывается принцип действия сети AlphaFold2. Данная сеть пытается предсказать 3D-структуру белка, основываясь на белковой последовательности. На вход сети подается последовательность белка. На первом шаге, с помощью специальных алгоритмов, производится поиск различных версий этого белка в специальной базе данных. Далее генерируются векторные эмбединги парного представления остатков аминокислот, на основе которых формируется 3D-структура белка. В данном методе используется подход Noisy Student Training - подход с частичным привлечением учителя. Так как для некоторых белков, уже известна 3D-структура, которая представлена в еще одной базе данных. Подход сначала обучает первую модель на размеченных данных, потом генерирует с помощью этой модели псевдо-разметку и тренирует уже другую модель на псевдо-разметке, совмещенной с правильной разметкой. При определении парного представления, используется более сложный механизм внимания, Он предполагает составления матрицы проекций по всем осям признакового пространства белков. Далее, нейросеть будет пытаться определить отношения между эмбедингами и несколькими парами таких матриц, после получения которых будет формироваться 3D-структура.