

Предсказание свойств и типов атомов в молекулярных графах при помощи сверточных сетей

Манучарян Вардан Аветисович

Московский физико-технический институт

Цель исследования

- Цель работы

Предложить метод предсказания типов атомов

- Проблема

Существующие модели предсказывают тип на основе гибридизации

- Метод решения

Использование свёрточных нейронных сетей, учитывающих окружение атома и 3d структуру молекулярных графов

Литература

- *Mathias Niepert, Mohamed Ahmed, Konstantin Kutzkov. Learning Convolutional Neural Networks for Graphs, 2015*
- *Steven Kearnes, Kevin McCloskey, Marc Berndl, Vijay Pande, Patrick Riley. Molecular Graph Convolutions: Moving Beyond Fingerprints, 2016*
- *Maria Kadukova, Sergei Grudinin. Knodle: A Support Vector Machines-Based Automatic Perception of Organic Molecules from 3D Coordinates // Journal of Chemical Information and Modeling, American Chemical Society, 2016, 56 (8), pp.1410-1419*

Постановка задачи

Пусть $\mathfrak{S} = \{\mathfrak{s}_1, \dots, \mathfrak{s}_m\}$ - множество атомов в различных молекулах. Пусть $\mathbf{y} = \{y_1, \dots, y_m\}$ - типы атомов.

Пусть $G = \{g_1, \dots, g_n\}$ - набор функций, таких что $\forall i \forall j$ g_j отображает \mathfrak{s}_i в (i, j) элемент матрицы X :

$$g_j : (b_j, \mathfrak{s}_i) \rightarrow x_{ij} \in \mathbb{R}^1,$$

где b_j - набор параметров для g_j .

Определим модель \mathbf{f} , сопоставляющую каждой строке X число из $[0, 1]$

$$f(w, X) = \frac{1}{1 + \exp(-Xw)},$$

где оптимальные параметры \hat{w} минимизируют функцию потерь

$$\hat{w} = \arg \min_w S(w|f, X, y),$$

где

$$S(w|f, X, y) = -\ln\left(\sum_{i=1}^m y_i \log f(x_i, w) + (1 - y_i) \log(1 - f(x_i, w))\right)$$

Цели эксперимента

- Продумать архитектуру сети, предсказывающей свойства вершин, зависящие от их окружения