

Классификация воображаемых движений

Данные прикреплены в письме (файл alex_long_1.mat). Данные представляют собой записи с 24 каналами. Классифицировать данные нужно на 3 группы: воображаемое движение левой рукой, воображаемое движение правой рукой и состояние покоя.

Для запуска классификации с помощью программного обеспечения Numenta (с помощью пакета nupic для языка Python 2.7), они были преобразованы в .csv файл определенного образца. Файл также прикреплен в письме (24chan_data_1.csv).

Частичное описание работы с платформой Numenta и программным пакетом nupic доступно по ссылке github.com/numenta/nupic/wiki.

Hierarchical Temporal Memory (HTM) — технология, "моделирующая" работу коры головного мозга. В основе этой технологии используется иерархическое представление отдельных "блоков памяти" (рис. 1a). Эти блоки (или регионы) организованы как трехмерный массив (рис. 1b). HTM регионы также как и области коры мозга используют разреженные распределенные репрезентации, то есть в регионе активен очень малый процент "нейронов" (отдельных ячеек) одновременно.

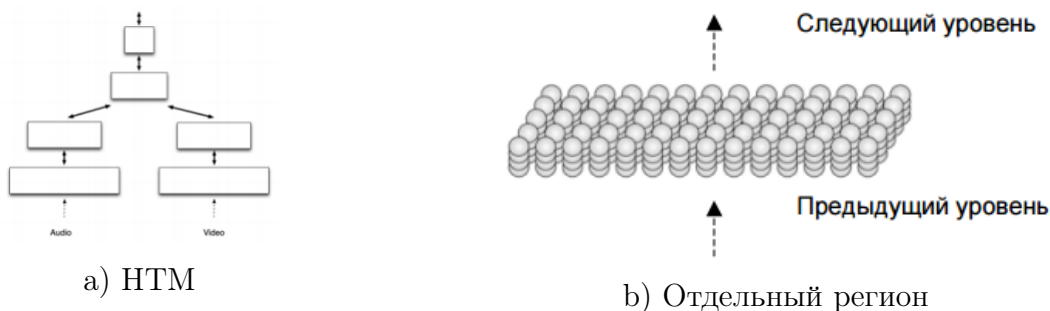


Рис. 1

Нижние уровни иерархии соответствуют Temporal Pooler, который отвечает за запоминание последовательностей, сильно изменяемых данных и контекстов для событий. Верхние слои соответствуют Spatial Pooler, отвечающий за обучение связей и представление более общих понятий (нежели последовательности).

Temporal Pooler используется для запоминания и предсказания на несколько шагов. В открытом коде nupic на github'е достаточно много примеров с применением пакета для предсказания некоторых данных, зависящих от времени. Для простой классификации данных используют Spatial Pooler.

Для обозначенных в начале данных была запущена классификация. Результаты работы скрипта приведены в документе 24chan_data_1_results.csv. Столбец classification отвечает за истинный класс, а столбец modelBestPrediction.0 — за результат классификации. Я не до конца интерпретировала результат, так как в части наблюдений в полученном значении появляется не номер класса, а строка вида:

3: 0.18060211411088037}

Но за исключением части не интерпретируемых результатов, классификация верная. Я найду путь решения проблемы, и после опишу результаты численно.

В целом, я пока подбирала параметры больше интуитивно. Далее подберу модель лучше.