Слайд 1

Добрый день, уважаемые члены государственной экзаменационной комиссии. Разрешите представить Вашему вниманию магистерскую диссертацию на тему «Разработка многоуровневой распределенной системы анализа больших данных биоподобными сенсорно-моторными алгоритмами с целью определения аномалий»

Слайд 2

Моя работа основывается на книге Джефа Хоккинса «1000 умов. Теория нового интеллекта» и статьях исследовательской группы Нумента, основателем которой так же является Джеф Хоккинс. Они занимаются исследованием неокортекса и биологических процессах, которые в нем протекают.

Слайд 3

Неокортекс – это область мозга, выделенная желтым цветом на рисунке А. Если мы возьмем его и развернем, то получим полотно 1х1 метр, состоящее из 5-6 слоев, слои можно увидеть на рисунке Б. При этом все слои будет пронзать такая структура, как кортикальная колонка, представленная на рисунке В. На этом же рисунке можно увидеть скопления точек в рамках одного слоя. Такие скопления представляют абстрактную структуру – мини-колонку. Точки же в свою очередь представляют нейроны, пример пирамидального нейрона представлен на рисунке Г.

Слайд 4.

Существующие нейронные сети используют терминологию биологии, но у них нет цели смоделировать процессы неокортекса. Они используют точечную модель нейрона. Нейрон теории 1000 мозгов сложнее, в нем моделируются проксимальные, базальные, апикальные зоны активности, а так же дендриты, синапсы, аксоны и различные состояния клеток.

Слайд 5.

Представленная мною тема является продолжением доклада прошлогодней конференции. Тогда была разработана однослойная биоподобная система, способная предсказать следующий символ в строке.

Слайд 6.

Основная цель работы – разработка универсального ядра системы, которое моделирует поведение одной кортикальной колонки, состоящей из двух слоев. И написание инструмента визуализации для отслеживания состояния системы и понимания процессов, протекающих в ядре.

Слайд 7.

Как я уже сказал, система является двухслойной, один слой отвечает за входные данные, второй – за выходные. Схема одного слоя представлена на рисунке Б. Слой состоит из мини-колонок (абстрактной структуры). Мини-колонки состоят из клеток, у каждой клетки есть множество дендритов, а дендриты содержат синапсы.

Слайд 8.

Так же я упомянул про сенсорно-моторные алгоритмы.

Это подразумевает то, что говоря о данных, мы будем знать, откуда они пришли и что они из себя представляют, то есть сенсорная информация. То, как кодируется эта информация, в текущей работе не рассматривалось, достаточно понимать, что сенсорный сигнал или сигнал местоположения представляет из себя разряженный набор битов, например десять единиц на две тысячи битов.

Перейдем к примеру, на слайде представлена общая модель системы.

Сперва в модуль адаптации, тот который кодирует данные поступает информация.

Затем на входной слой поступает сигнал местоположения (синяя стрелочка). Этот сигнал деполяризует клетки входного слоя, это образует некоторый заряд на клетке. То есть эта клетка готовится стать активной.

Теперь посылаем сенсорный сигнал, который в первую очередь активирует клетки, которые были деполяризованы.

В это же время, если до этого на выходном слое была некоторая активность, то она может деполяризовать клетки выходного слоя (боковые синие стрелочки).

Теперь входной слой посылает сигнал на выходной слой, активируя его клетки.

После активации происходит отправка данных для декодирования информации, чтобы проверить, а не удалось ли нам определиться с предсказанием.

А так же происходит фидбек от выходного слоя к входному. То есть еще один процесс предсказания следующей активности входного слоя.

Слайд 9.

На самом деле, когда мы говорим о биоподобной системе, то подразумеваем, что у нас есть несколько датчиков, иными словами - сенсоров. И на каждый сенсор есть несколько колонок. И колонки, находящиеся на границе областей, могут обмениваться информацией.

Слайд 10.

С биологической точки зрения это объясняется тем, что за обработку информации с различных частей тела у нас отвечают целые области неокортекса, которые тесно соседствуют между собой. Это видно на изображении.

Слайд 11.

На текущем слайде представлена математическая модель. Здесь приведены формулы для деполяризации клеток входного слоя, активации клеток входного и выходного слоя, а так же для обучения клеток, на примере дендритных сегментов.

Слайд 12.

Ядро системы реализовано на функциональном языке программирования Erlang. Такой выбор сделан потому, что erlang хорошо работает с легковесными процессами, что позволяет с минимальными затратами реализовать распределенную систему.

Модуль визуализации написан на .Net WPF с применением паттерна MVVM.

Слайд 13.

Для обмена информацией между модулями используются TCP сокеты. Был написан небольшой синтаксический анализатор, для формирования пакетов данных.

Слайд 14.

На текущем слайде представлен интерфейс программы. Здесь можно увидеть состояние клеток входного и выходного слоя.

По каждой клетке входного слоя можно увидеть активность дендритов и синапсов, а так же увидеть дендрит, который привел к активации клетки.

Можно проанализировать связи между слоями.

Функционал позволяет отправлять сигналы местоположения и сенсорные сигналы в ядро системы, а так же выгружать дамп состояния.

Слайд 15.

В статье Хоккинса приводится анализ емкости системы. Сравнивается влияние количества мини-колонок и клеток выходного слоя на количество объектов, которые может запомнить сеть при 10 ощущениях на объект.

Таким образом сеть из 150 мини-колонок и четырех тысяч клеток выходного слоя способна запомнить порядка 400 объектов.

Слайд 16.

Выявление аномалий выглядит следующим образом. Ядро системы делает предсказание на основе данных хт и выводит паттерн активации. Вычисляется ошибка прогнозирования – то, на сколько паттерн активации отличается от паттерна прогнозирования. И на основе ошибки вычисляется вероятность того, в данных произошла аномалия.

Слайд 17.

В ходе работы была достигнута основная цель – реализовано ядро системы и модуль визуализации.

Слайд 18.

Следующие цели – изучение новых статей группы Нумента, понимание того, как мозг интерпретирует сигналы местоположения, изучение влияние других типов клеток, таких как grid cells и displacement cells, и написание модуля адаптации данных для ядра системы.