

М. Г. Киселева – студентка кафедры компьютерной математики и программирования

М. Д. Поляк – научный руководитель

МОДЕЛИРОВАНИЕ СЛОЖНЫХ ПРОЦЕССОВ С ПОМОЩЬЮ МАШИНЫ БОЛЬЦМАНА

Предложен нейросетевой подход к решению проблемы моделирования сложных процессов на основе структуры, называемой Машиной Больцмана.

Цель исследования состоит в изучении нейронной сети – Машины Больцмана – при моделировании сложных процессов.

Объектом исследования являются искусственные нейронные сети, а также методы нейросетевого моделирования.

Машина Больцмана представляет собой хороший инструмент для моделирования данных. Использование стохастической нейронной сети – Машины Больцмана – позволяет построить модель, описывающую распределение исходных данных.

Машина Больцмана представляет собой стохастическую машину, компонентами которой являются стохастические нейроны. Стохастический нейрон находится в одном из двух вероятных состояний. Этим двум состояниям можно формально присвоить +1 (соответствующее включенному состоянию) и -1 (соответствующее выключенному состоянию). Аналогично можно принять значениями этих состояний 1 и 0 соответственно. Стохастические нейроны Машины Больцмана разбиваются на две функциональные группы: видимые и скрытые. Видимые нейроны предоставляют интерфейс между сетью и средой, в которой она работает. Скрытые нейроны работают свободно – они используются для выражения ограничений, содержащихся во входных векторах*.

Обучение Машины Больцмана – весьма длительный и трудоемкий процесс, но оно может быть достаточно эффективно реализовано в архитектуре, называемой Машиной Больцмана с ограничениями, в которой отсутствуют связи между нейронами внутри слоя (рис.1).

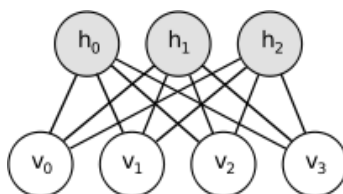


Рис.1. Ограниченная Машина Больцмана

Машину Больцмана с ограничениями также иногда называют моделью на основе произведения экспертов. В такой модели каждый нейрон скрытого слоя отвечает за отдельный признак.

В качестве моделируемых данных рассматриваются искусственные данные. Для обучения сети используется набор искусственных данных, представляющих собой 500 точек, формирующих линейно неразделимое множество. Данные подчиняются равномерному закону распределения (рис.2).

* Хайкин С. Нейронные сети: полный курс, 2-е издание. : Пер. с англ. – М. : Издательский дом «Вильямс», 2006.

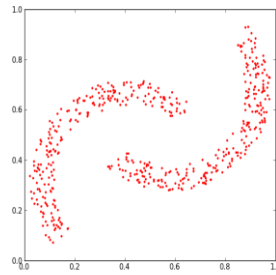


Рис. 2. Искусственные данные

Структура сети в данном случае выглядит следующим образом (рис. 3). В видимом слое содержится 2 нейрона, 8 нейронов – в скрытом слое. Количество эпох обучения взято равным 4000.

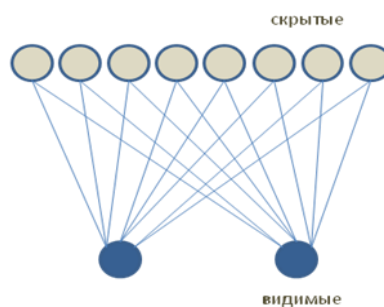
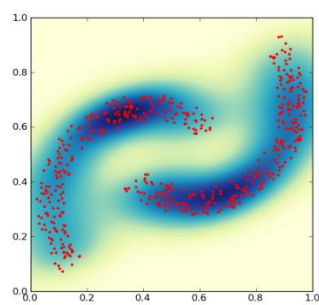
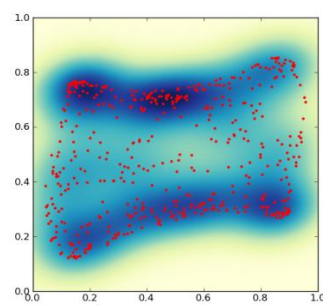


Рис. 3 Структура сети

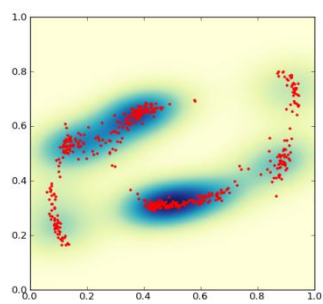
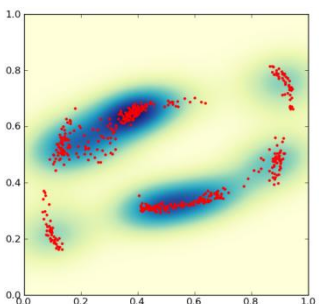
После проведения моделирования были получены результаты, являющиеся реконструкцией исходных данных (рис. 4,а). На вход Машины Больцмана подавались случайные точки, равномерно распределенные в квадрате от 0 до 1 по обеим осям. Реконструированные данные представляют собой данные, воспроизводимые Машиной Больцмана на первом (рис. 4,б), пятидесятом (рис. 4,в) и сотом (рис. 4,г) шаге цепи Маркова.



а)



б)



в)

г)

Рис. 4.

а) исходные данные; б) реконструированные данные на 1 шаге;
в) реконструированные данные на 50 шаге; г) реконструированные данные на 100 шаге

В ходе исследования было проведено моделирование, где были установленные параметры и изменяющийся параметр. Фиксированные параметры: число видимых нейронов взято равным 2, количество эпох обучения – 4000. Изменяющиеся параметры: число нейронов в скрытом слое – 6 (рис. 5,а), 10 (рис. 5,б) и 12 (рис. 5,в). Реконструированные данные взяты на пятидесятом шаге цепи Маркова.

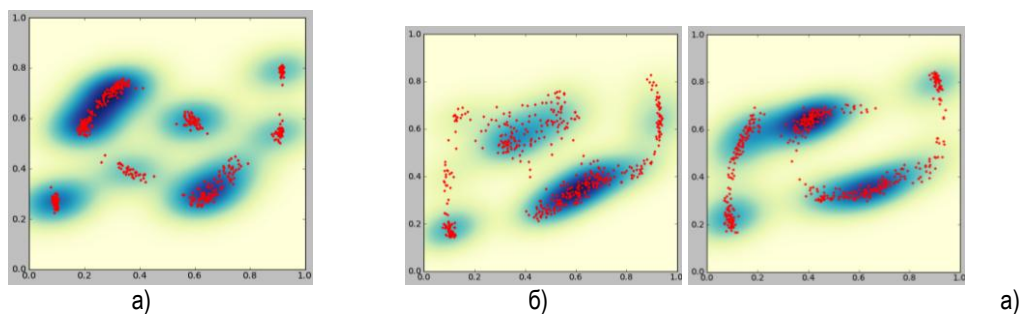


Рис. 5.

а) реконструированные данные при 6 нейронах; б) реконструированные данные при 10 нейронах;
в) реконструированные данные при 12 нейронах

Очевидно, что с увеличением числа нейронов в скрытом слое результат моделирования улучшается. Но даже при числе скрытых нейронов, равным 12 (рис. 5,а), он хуже, чем при 8 скрытых нейронах (рис. 4,в).