



О.П. Солдатова, А.Ю. Разинов

ПРИМЕНЕНИЕ ЭВОЛЮЦИОННЫХ МЕТОДОВ ДЛЯ ГЕНЕРАЦИИ ОПТИМАЛЬНОЙ СТРУКТУРЫ НЕЙРОННЫХ СЕТЕЙ В ПРОЦЕССЕ ОБУЧЕНИЯ

(ФГБОУ ВПО «Самарский государственный аэрокосмический университет им. академика С.П. Королева (национальный исследовательский университет)»)

В рамках данного доклада рассматривается эволюционный подход для генерации оптимальной структуры нейросети в процессе обучения. Эволюционные принципы используются при решении различных вычислительных задач, таких как задачи оптимизации и моделирования. Выделяют следующие основные виды эволюционных алгоритмов: генетические алгоритмы; эволюционное программирование; эволюционные стратегии; генетическое программирование [1]. В докладе представлены несколько реализаций эволюционных методов, являющихся частным случаем генетического алгоритма, примененных для модификации существующих алгоритмов обучения нейронных сетей с целью генерации оптимальной для данной задачи структуры сети.

Для обучения нейронных сетей применяются алгоритмы обучения с учителем или без учителя. Большинство из применяемых алгоритмов обучения являются итерационными, и обучение нейросетей сводится к следующим этапам:

1. Генерация начального приближения параметров сети.
2. Подача на вход сети обучающего примера и расчёт выходных значений сети.
3. Расчёт функции ошибки и сравнение её с заданным порогом.
4. Если порог ошибки или предельное число итераций не достигнуто, то корректировка параметров сети.
5. Иначе – завершение обучения.

Алгоритм обучения предполагает убывающее значение нахождения глобального минимума функции ошибки. Если обучение не привело к желаемому результату, обучение проводится заново с измененными параметрами обучения.

При реализации классических алгоритмов обучения нейронных сетей типичными проблемами являются:

1. Преждевременное схождение в локальный оптимум.
2. Выбор коэффициента обучения.
3. Определение оптимальных параметров сети.

Все эти проблемы можно решить с помощью эволюционных методов. Проблема преждевременного схождения решается благодаря учету перспективности решения. Проблема выбора коэффициента обучения решается автоматической подстройкой коэффициента для каждого решения на каждой итерации эволюционного алгоритма.

Эволюционное моделирование или эволюционное вычисление (evolutionary computation) — класс методов поиска решений и глобальной оп-



тимизации, основанный на итеративной адаптации под заданные условия, такие как ограничения и целевые функции. Эволюционные вычисления делятся на две независимые категории – эволюционные алгоритмы и «роевой интеллект». В основе эволюционного моделирования лежат реальные биологические эволюционные процессы — поведение популяции или сообществ животных в эволюционных алгоритмах или поведения одного сообщества живых организмов или группы неодушевленных предметов в методах роевого интеллекта. Эволюционное моделирование используется как средство глобальной оптимизации, подбора параметров алгоритмов, автоматического проектирования.

Обычно эволюционные методы используются как самостоятельные методы оптимизации. Существует отдельный термин для эволюционных методов, примененных для оптимизации структуры нейросетей – нейроэволюция [2]. При нейроэволюции эволюционным алгоритмом подбираются веса связей между нейронами или другие параметры нейросети. В данном докладе рассматривается эволюционный алгоритм, примененный для обучения нейронных сетей, который не является самостоятельным методом оптимизации и не связан с термином нейроэволюция. Эволюционный алгоритм здесь является средством улучшения работы классического метода обучения. Основная особенность описываемого метода – это простота реализации при достаточно заметном улучшении работы классического алгоритма. Данный метод не описан в литературе, однако является наследником классического генетического метода.

Данный алгоритм очень похож на классический алгоритм обучения нейросети. Он состоит из следующих шагов:

1. Генерация начального приближения параметров N сетей. ($N > 1$)
2. Обучение каждой сети по классическому алгоритму в течение P итераций.
3. Отбор M лучших нейросетей на основе целевой функции, отбрасывание худших ($N-M$), копирование лучших нейросетей на место худших. ($M < N$)
4. Если порог ошибки лучшей нейросети или предельное число итераций не достигнуто, случайно изменить коэффициент обучения и численные параметры каждой нейросети, повторить алгоритм с шага 2.
5. Иначе – завершение обучения.

Суть данного алгоритма – отбрасывание худших решений в процессе обучения. Данный алгоритм автоматически подбирает коэффициент обучения, причем, решения с неподходящим коэффициентом обучения отбрасываются на этапе отбора. Алгоритм имеет возможность выхода из локального оптимума благодаря случайному поиску и возможности изменять направление поиска.

На практике данный алгоритм при увеличении времени вычисления пропорционально количеству нейросетей в группе не дает существенного преимущества в силу ряда недостатков: перспективные решения, не являющиеся лучшими в группе на текущей итерации отбора могут быть отброшены в пользу лучших в данный момент, но неперспективных в будущем. Также в связи с



этим затруднен выход из локального оптимума. Таким образом, недостатки алгоритма сводятся к тому, что не учитывается перспективность решений.

Развивая представленный метод поиска, можно получить эволюционный алгоритм с учетом перспективы. Практика показала, что такой подход оправдывает дополнительные расходы времени, в отличие от изложенного ранее простого эволюционного алгоритма, и позволяет получать решения более близкие к глобальному оптимуму. Именно этому методу посвящен доклад. При обучении нейросетей метод способен выделить наиболее перспективные варианты решений и отсеять решения, которые в ближайшем будущем могут привести к тупиковой ситуации.

Суть данного метода заключается в оценке не настоящего состояния решений, а будущего. Будущее оценивается перед каждой операцией отбора и его оценка является целевой функцией описываемого метода оптимизации. Также как и раньше внутри каждой итерации алгоритма каждое решение, то есть каждая нейросеть, развивается, обучается по классическому алгоритму вроде алгоритма обратного распространения ошибки. Нововведение заключается в дальнейшем продолжении обучения по классическому алгоритму, то есть после точки, являющейся настоящим, моделируется будущее развитие решения. По окончании моделирования высчитывается общее улучшение работы нейросети, как для фазы обычного обучения, так и для моделирования будущего. Улучшение пропорционально уменьшению ошибки нейросети на обучающей выборке. В конце фазы моделирования будущего решение откатывается к точке, являющейся границей первой и второй фазы. То есть моделирование будущего не изменяет решения и используется только для подсчета целевой функции. Если бы изменения не терялись, вторая фаза не была бы моделированием будущего и тупиковые решения не могли бы быть определены.

В результате, имея оценку развития каждого решения на этапе отбора, можно отбросить тупиковые решения до того, как вся группа решений зайдёт в тупик. Практика показала эффективность метода. При использовании описанной сравнительно простой надстройки над классическим методом обучения улучшается качество обучения нейросети, то есть достижение заданного значения ошибки в результате обучения за один запуск программы. В представленном алгоритме попадание в тупик предсказывается, перезапуск алгоритма не требуется, так как алгоритм способен выйти из оптимума самостоятельно.

На рисунке 1 представлены графики зависимостей ошибки предсказания радиально-базисной сетью котировок биржевых акций евро-доллар от времени обучения (от 0 до 1000 секунд) при использовании эволюционного метода и градиентного спуска (слева) и при использовании только метода градиентного спуска. Из графика видно, что метод градиентного спуска быстро справился с задачей, однако в перспективе решение, сделанное с помощью эволюционного метода оказалось лучше. В ходе эксперимента установлено двукратное уменьшение ошибки предсказания при использовании эволюционного метода с учетом перспективности.

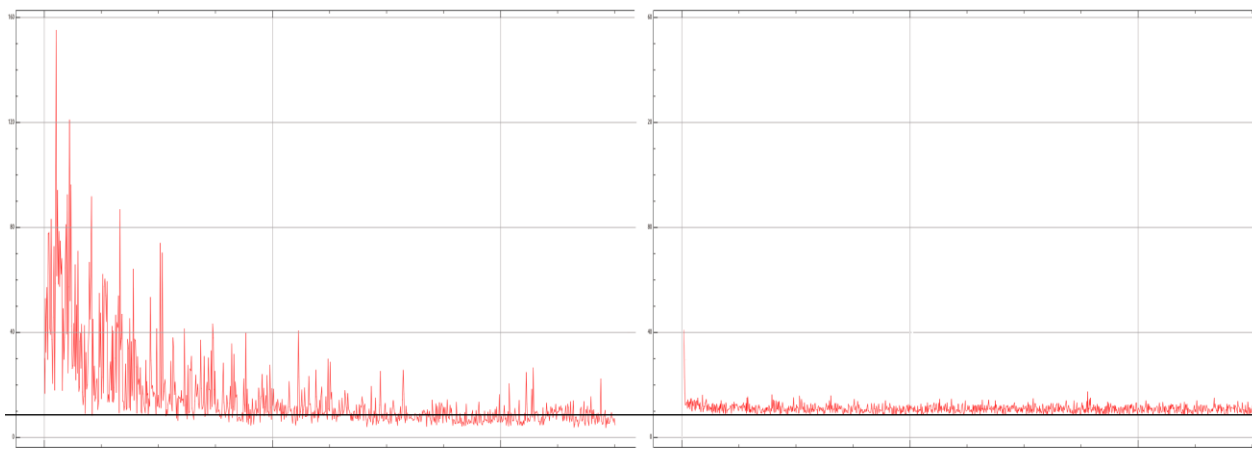


Рис. 1. Графики зависимости ошибки предсказания от времени

Исходные данные эксперимента:

- биржевые данные взяты за 4 года, содержат 39 млн. значений цены акций;
- время обучения – 1000 секунд;
- при вычислении ошибки предсказания использовались 500 случайных отрезков истории курса валют в качестве входных и выходных данных, максимальная величина ошибки – 500, минимальная – 0;
- количество нейронов во входном слое – 10;
- количество нейронов в скрытом слое – 10;
- длина фазы моделирования будущего развития в эволюционном методе – 90% длины всей итерации;
- количество нейросетей в конкурирующей группе – 100.

Литература

1. Цой Ю.Р., Спицын В.Г. Эволюционный подход к настройке и обучению искусственных нейронных сетей // Нейроинформатика, 2006. – том 1. – №1. – с. 35 – 61
2. Gomez F., Miikkulainen R. Learning robust non-linear control with neuroevolution. Technical Report AI01-292. – Department of Computer Sciences, The University of Texas at Austin, 2001. – pp. 1 – 4