

Нейротехнология обработки больших данных с использованием векторного критерия

П. В. Четырбок

ФГАОУ ВО «Крымский федеральный университет им. В.И. Вернадского»

Аннотация. Рассматриваются особенности применения нейротехнологии для обработки больших данных. Описаны типичные характеристики больших данных. Показаны проблемы, возникающие при использовании нейротехнологии для обработки больших данных. Представлены методы решения проблем, включающие использование специальных типов нейронных сетей (НС), редукцию к задачам меньшей размерности и декомпозицию данных с использованием векторного критерия.

Ключевые слова: *большие данные; нейронные сети; декомпозиция данных; векторный критерий*

I. ПОСТАНОВКА ПРОБЛЕМЫ

К области нейротехнологии относится направление, связанное с обработкой информации на основе принципов функционирования естественных нейронных систем. Математическое моделирование таких систем привело к созданию искусственных нейронных сетей. В биологических сетях принципиальное значение имеет структура связей между нервными клетками, называемых нейронами. Обособленные группы нейронов в нейробиологии называются нейронными модулями. Они взаимодействуют между собой только через внешние рецепторные и аксоновые поля. Через скрытые нейроны модулей идет обработка информации, которая поступает на рецепторы, и формируется реакция в выходном аксоновом поле модуля. Если предположить, что нейронные модули в сети не имеют пересечений рецепторных или аксоновых полей, то получим модель модульной нейронной сети с инъективными связями. Инъективность связей позволяет представить структуру нейронной сети в виде ориентированного графа. Погружение структурной модели в пространство рецепторных и аксоновых полей порождает топологическую модель нейронной сети, по которой можно реализовать адаптивный алгоритм обработки данных [1]. Особенно такая обработка актуальна для больших данных. Она позволяет увеличить скорость обработки в реальном режиме и уменьшить объем обрабатываемых данных для классификации.

II. ЦЕЛЬ СТАТЬИ

Рассматриваются возможности оптимальной модульной нейронной сети, топологическая модель которой построена с использованием не только пространств рецепторных и аксоновых полей, но и пространства ошибок, полученного с помощью векторного критерия, для обработки больших данных (Data mining technology).

III. ОСНОВНАЯ ЧАСТЬ

Векторный критерий позволяет построить оптимальную модульную нейронную сеть для распознавания объектов, исключив из нее модули которые не влияют на процесс распознавания [2]. То есть в пространстве рецепторов исключаются поля рецепторов не оказывающие существенное влияние на распознавание объектов по определенному правилу.

Правило исключения модуля из сети следующее:

- если векторный критерий при распознавании нового объекта на модуле не изменяет своего значения, то этот модуль может быть безболезненно исключен из модульной нейронной сети.

Применяя данное правило, мы фактически строим оптимальную модульную нейронную сеть для распознавания нового объекта. В конкретной реализации нейронной сети с каждым нейронным модулем связаны два линейных векторных пространства:

- пространства рецепторов;
- пространство аксонов.

Добавим к ним еще пространство ошибок. Пространство ошибок содержит вектора ошибок, полученные при распознавании объектов нейронным модулем.

Для нейронного модуля размерность пространства рецепторов равна числу рецепторов, а размерность пространства аксонов равна числу аксонов. То есть нейронный модуль является оператором, который преобразует вектор из пространства рецепторов в пространство аксонов. Оператор можно задать, вводя базисы в векторном пространстве.

Если за базис взять вектора – типичные представители классов объектов из обучающей выборки. Типичный представитель класса объектов выбирается из обучающей выборки среди эталонных образцов объектов этого класса по векторному критерию. А именно выбирается тот эталон, на котором векторный критерий максимален. Векторный критерий позволяет в обучающей выборке среди эталонных образцов выбрать опорные вектора для машины опорных векторов. Также мы сможем добавить еще один базис, а именно вектора, полученные на ошибках распознавания опорных векторов в пространстве ошибок. Исходя из этого базиса, мы сможем вычислить векторный критерий в пространстве ошибок при распознавании нового объекта, и тем самым используя вышеприведенное правило оптимизировать модульную нейронную сеть.

Минимизируемой целевой функцией ошибки модуля НС является величина [3]:

$$E(w) = \frac{1}{2} \sum_{j,p} (y_{j,p}^{(N)} - d_{j,p})^2 \quad (1)$$

где $y_{j,p}^{(N)}$ – реальное выходное состояние нейрона j выходного слоя N модуля нейронной сети при подаче на ее входы p -го образа; $d_{j,p}$ – идеальное (желаемое) выходное состояние этого нейрона.

Суммирование ведется по всем нейронам выходного слоя и по всем обрабатываемым сетью образам. Минимизация ведется методом градиентного спуска, что означает подстройку весовых коэффициентов следующим образом:

$$\Delta w_{ij}^{(n)} = -\eta \cdot \frac{\partial E}{\partial w_{ij}} \quad (2)$$

Здесь w_{ij} – весовой коэффициент синаптической связи, соединяющей i -ый нейрон слоя $n-1$ с j -ым нейроном слоя n , η – коэффициент скорости обучения, $0 < \eta < 1$.

$$\frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial y_j} \cdot \frac{dy_j}{ds_j} \cdot \frac{\partial s_j}{\partial w_{ij}} \quad (3)$$

Здесь под y_j , как и раньше, подразумевается выход нейрона j , а под s_j – взвешенная сумма его входных сигналов, то есть аргумент активационной функции. Так как множитель dy_j/ds_j является производной этой функции по ее аргументу, из этого следует, что производная активационной функции должна быть определена на всей оси абсцисс. В связи с этим функция единичного скачка и прочие активационные функции с неоднородностями не подходят для рассматриваемых НС. В них применяются такие гладкие функции, как гиперболический тангенс или классический сигмоид с экспонентой. В случае гиперболического тангенса

$$\frac{dy}{ds} = 1 - s^2 \quad (4)$$

Третий множитель $\partial s_j / \partial w_{ij}^{(n)}$ очевидно, равен выходу нейрона предыдущего слоя $y_i^{(n-1)}$.

Что касается первого множителя в (3), он раскладывается следующим образом:

$$\frac{\partial E}{\partial y_j} = \sum_k \frac{\partial E}{\partial y_k} \cdot \frac{dy_k}{ds_k} \cdot \frac{\partial s_k}{\partial y_j} = \sum_k \frac{\partial E}{\partial y_k} \cdot \frac{dy_k}{ds_k} \cdot w_{jk}^{(n+1)} \quad (5)$$

Здесь суммирование по k выполняется среди нейронов слоя $n+1$.

Введя новую переменную

$$\delta_j^{(n)} = \frac{\partial E}{\partial y_j} \cdot \frac{dy_j}{ds_j} \quad (6)$$

мы получим рекурсивную формулу для расчетов величин $\delta_j^{(n)}$ слоя n из величин $\delta_k^{(n+1)}$ более старшего слоя $n+1$.

$$d_j^{(n)} = \left[\sum_k d_k^{(n+1)} \cdot w_{jk}^{(n+1)} \right] \cdot \frac{dy_j}{ds_j} \quad (7)$$

Для выходного же слоя

$$d_i^{(N)} = (y_i^{(N)} - d_i) \cdot \frac{dy_i}{ds_i} \quad (8)$$

Теперь мы можем записать (2) в раскрытом виде:

$$Dw_{ij}^{(n)} = -h \cdot d_j^{(n)} \cdot y_i^{(n-1)} \quad (9)$$

Иногда для придания процессу коррекции весов некоторой инерционности, сглаживающей резкие скачки при перемещении по поверхности целевой функции, (9) дополняется значением изменения веса на предыдущей итерации

$$Dw_{ij}^{(n)}(t) = -h \cdot (m \cdot Dw_{ij}^{(n)}(t-1) + (1-m) \cdot d_j^{(n)} \cdot y_i^{(n-1)}) \quad (10)$$

где μ – коэффициент инерционности, t – номер текущей итерации.

Таким образом, полный алгоритм обучения НС с помощью процедуры обратного распространения строится так:

1. Подать на входы модуля сети один из возможных образов и в режиме обычного функционирования модуля НС, когда сигналы распространяются от входов к выходам, рассчитать значения последних. Напомним, что

$$s_j^{(n)} = \sum_{i=0}^M y_i^{(n-1)} \cdot w_{ij}^{(n)} \quad (11)$$

где M – число нейронов в слое $n-1$ с учетом нейрона с постоянным выходным состоянием $+1$, задающего смещение; $y_i^{(n-1)} = x_{ij}^{(n)}$ – i -ый вход нейрона j слоя n .

$$y_j^{(n)} = f(s_j^{(n)}), \text{ где } f() \text{ – сигмоид} \quad (12)$$

$$y_q^{(0)} = I_q, \quad (13)$$

где I_q – q -ая компонента вектора входного образа.

2. Рассчитать $\delta^{(N)}$ для выходного слоя по формуле (8).

Рассчитать по формуле (9) или (10) изменения весов $\Delta w^{(N)}$ слоя N .

3. Рассчитать по формулам (7) и (9) (или (7) и (10)) соответственно $\delta^{(n)}$ и $\Delta w^{(n)}$ для всех остальных слоев, $n=N-1, \dots, 1$.

4. Скорректировать все веса в модуле НС

$$w_{ij}^{(n)}(t) = w_{ij}^{(n)}(t-1) + Dw_{ij}^{(n)}(t) \quad (14)$$

5. Если ошибка модуля сети существенна, перейти на шаг 1. В противном случае – конец.

Алгоритм нахождения вектора ошибок произвольного объекта описан в [4].

Модульная нейронная сеть обладает морфологией, представленной совокупностью нейронных модулей и межмодульных связей, поэтому использование теории категорий для построения математических моделей отвечает

природе модульной сети. Математическая категория как модель в целом представляет все возможные реализации нейронной сети заданной структуры, т.е. описывает весь класс проектных решений. В зависимости от способа построения категория нейронной сети может описывать либо уровень обработки данных, либо уровень топологии межслойных связей, либо уровень структуры модульной сети. Математическое моделирование позволяет упростить задачи анализа и синтеза нейронных сетей. В статье рассматривается категория как модель уровня обработки данных.

Рассмотрим категорию R векторных пространств. Объектами категории являются конечномерные пространства (рецепторов, аксонов и ошибок). Категория R – это набор объектов (ObR) вместе с набором отображений между объектами ($MorR$). Отображения называются морфизмами категории и обладают следующими свойствами:

1. каждый морфизм категории R принадлежит только одной паре объектов $A, B \in ObR$ (так как построенное нами отображение является отношением эквивалентности, то данное свойство выполняется);
2. в классе $MorR$ введена частичная бинарная операция умножения: произведение $\alpha\beta$ морфизмов $\alpha \in Mor(A, B)$ и $\beta \in Mor(C, D)$ определено тогда и только тогда, когда объект B совпадает с объектом C , и в этом случае $\alpha\beta \in Mor(A, D)$ (так как построенные нами отображения являются отношением эквивалентности, то данное свойство выполняется);
3. частичное умножение ассоциативно: $(\alpha\beta)\gamma = \alpha(\beta\gamma)$ для любых трех морфизмов $\alpha \in Mor(A, B)$, $\beta \in Mor(B, C)$, $\gamma \in Mor(C, D)$ (так как построенные нами отображения являются отношениями эквивалентности, то данное свойство выполняется);
4. в каждом множестве $Mor(B, B)$ содержится такой морфизм 1_B , называемый единичным морфизмом объекта B , что $\alpha 1_B = \alpha$ и $1_B \beta = \beta$ для любых морфизмов $\alpha \in Mor(A, B)$ и $\beta \in Mor(D, C)$ (так как построенные нами отображения являются отношениями эквивалентности, то данное свойство выполняется).

Для категории конечномерных пространств переход к двойственной категории R^* реализуется сопоставлением

каждому линейному пространству A сопряженного ему пространства A^* . Для построенной категории пространство ошибок будет являться сопряженным к пространству рецепторов. Категория R^{**} – двойственная категории R – отождествляется с категорией R . Операцию «*» можно рассматривать как отображение между категориями удовлетворяющее условиям:

1. $A^{**} = A$ для любого объекта категории R ;
2. $\alpha^{**} = \alpha$ для любого морфизма категории R .

То есть прямую и двойственную категорию можно рассматривать как одну категорию. Следовательно, категории пространства рецепторов и пространства ошибок можно рассматривать как одну категорию.

IV. Выводы

Оптимальная модульная нейронная сеть с векторным критерием позволяет построить алгоритм обработки больших данных, который существенно сокращает их объем и тем самым увеличивает скорость их обработки. Сокращение объема происходит за счет исключения из оптимальной модульной нейронной сети модулей, а заодно и рецепторных полей, не оказывающих существенного влияния на распознавание новых объектов. Векторный критерий распознавания объектов позволяет построить оптимальную модульную нейронную сеть для системной декомпозиции больших данных на однородные области малой размерности, для каждого модуля которой можно построить частный классификатор со своей областью компетенции.

СПИСОК ЛИТЕРАТУРЫ

- [1] Дорогов А.Ю. Теория и проектирование быстрых перестраиваемых преобразований и слабосвязанных нейронных сетей. –СПб.: «Политехника», 2014. 328 с.
- [2] Четырбок П.В. Построение решающего правила для классификации образов на основе векторов ошибок. Київ.: «Системні дослідження та інформаційні технології», 2013. №2. С. 114–120.
- [3] Саймон Нейронные сети: полный курс, 2-е издание: Пер. с англ. М.: Издательский дом «Вильямс», 2006. 1104 с
- [4] Chetyrbok P.V. Preliminary systemic decomposition of big data for their classification using vector criteria dynamic management model of innovations generations 2017 XX IEEE International Conference on Soft Computing and Measurements (SCM), 2017. Pages: 762–764