УДК 519.7:681.3

В.И. Васильев, С.Н. Эш

Особенности алгоритмов самообучения и кластеризации

Описан метод автоматической классификации, не требующий предварительного задания числа образов (классов), основанный только на анализе расположения точек смешанной обучающей выборки в выбранном пространстве. Приведен пример применения предложенного метода для решения задач самообучения и кластеризации.

A method is described of automatic classification without preliminary assignment of number forms (classes), based only on the analysis of the arrangement of the points of the mixed instruction election in the chosen space. An example of the application of the suggested method for solving the problem of self-education and clasterization is presented.

Описано метод автоматичної класифікації, який не потребує попереднього завдання числа образів (класів), основою якого є тільки аналіз розташування точок змішаної навчальної вибірки у вибраному просторі. Наведено приклад застосування запропонованого методу для розв'язання задач самонавчання і кластеризації.

Введение. С появлением первых систем, используемых для обучения распознаванию образов, появились очень заманчивые идеи о возможности построения алгоритмов, способных к самообучению, т.е. к самопроизвольному разделению изображений на классы без помощи учителя. В качестве изображения объекта понимается вектор, составляющие которого есть свойства объектов. Такие идеи были естественным откликом на появление распознающих систем перцептронного типа [1, 2], так как одна из возможных схем перцептрона (перцептрон с *R*-управляемым подкреплением) предполагала такой процесс обучения, при котором подкрепление связей всегда постоянно, а его знак полностью определяется текущей реакцией перцептрона, т.е. сигналом *R*-элемента, которым он реагирует на входное изображение (величина и знак подкрепления не зависят от принадлежности входного изображения тому или иному образу, а поэтому и не требуется указаний учителя об этой принадлежности).

Вопрос о принципиальной возможности самообучения был предметом длительных дискуссий. Долгое время было не ясно, что же такое самообучение, а поэтому возможность его осуществления оставалась сомнительной. Казалась бессмысленной, например, постановка задачи, при которой система должна угадать априорную классификацию учителя. Если же потребовать, чтобы система проводила какую-нибудь классификацию, то задача становится тривиальной, так как любое логическое устройство проводит классификацию входных сигналов.

Впоследствии оказалось, что, с одной стороны, по некоторым признакам можно угадать классификацию учителя в результате изучения изображений в заданном им пространстве, а с другой – задача классификации, зависящей от заданного исходного пространства, не так уж тривиальна.

Доверие и интерес к проблеме самообучения вновь вызвали работы [3–5], в которых четко поставлена задача самообучения, а также выяснены возможности систем перцептронного типа. Однако не следует забывать, что первые работы, в которых была затронута проблема самообучения [1–2,], сыграли положительную роль, так как они продолжали привлекать интерес к проблеме во времена глубокого пессимизма. Кроме того, авторы этих работ были не так уж далеки от современной формулировки проблемы.

Свойство разделимости

Если бы удалось подметить всеобщее свойство, которое не зависит от природы образов, а определяет лишь их способность к разделимости, то наряду с обычной задачей обучения распознаванию с использованием информации о принадлежности каждого объекта наблюдения из обучающей последовательности тому или иному образу, можно было бы поставить другую классификационную задачу – обучения без учителя. Задачу такого рода на описательном уровне можно поставить следующим образом: одновременно или последовательно предъявляются объекты без каких-либо указаний учителя о принадлежности этих объектов образам. Входное устройство воспринимает множество объектов и, используя некоторое свойство разделимости образов, проводит классификацию объектов. После такого процесса приобретается способность к классификации объектов не только из обучающей последовательности, но и тех, которые ранее не предъявлялись. Процессом самообучения системы называют процесс, в результате которого эта система без указаний учителя приобретает способность к выработке одинаковых реакций на изображения объектов одного и того же образа и различных реакций на изображения различных образов. Роль учителя при этом состоит лишь в задании системе некоторого объективного свойства, одинакового для всех образов, определяющего способность к разделению множества объектов на образы.

Может показаться, что такое свойство задать невозможно хотя бы потому, что классифицируемые объекты обладают различными признаками и система не может сама решать, какие из них принимать во внимание, а какие отбрасывать, так как она не может разгадать априорную классификацию, задуманную учителем. Но если предположить, что учитель сам задает пространство признаков, исходя из задуманной классификации, то положение становится более обнадеживающим. Очевидно, что, указывая пространство признаков, учитель тем самым задает то свойство, определяющее разделимость образов, которое достаточно системе, чтобы оказаться способной к самообучению. Действительно, любая классификация, производимая человеком, характеризуется теми параметрами, по которым находится не всегда точно определенное сходство объектов, подлежащих классификации. Следовательно, для решения задачи самообучения прежде всего необходимо формализовать интуитивное понятие сходства.

Рассмотрим наиболее распространенную интерпретацию сходства, основанную на геометрических представлениях. Введем на множестве X, каждому элементу которого соответствует единственное изображение, некоторую метрику, определяющую пространство классификации. Тогда два объекта тем более похожи друг на друга, чем ближе соответствующие им точки в смысле метрики, введенной на множестве X. Целесообразность такого понимания сходства полностью зависит от выбранной метрики: если она неадекватна решаемой задаче, то понятие сходства не будет адекватным тому интуитивному понятию, которое подразумевается учителем и определяет классификацию объектов. Отождествление интуитивного понятия сходства объектов с близостью соответствующих им точек в пространстве, определяемом некоторой системой признаков, заранее указанной учителем, позволило предположить, что классу схожих объектов (т.е. образу) соответствует компактное множество точек, а разным классам (различным образам) соответствуют удаленные друг от друга компактные множества.

Взаимное расположение точек уже содержит информацию о том, как следует разделить множество точек, соответствующих объектам в пространстве, заданном учителем. Эта информация и определяет то свойство разделимости образов, которое оказывается достаточным для самообучения системы распознаванию образов. Выбор метрики или системы признаков для каждой конкретной задачи – творческий процесс, и задача его формализации в статье не рассматривается.

Разделимость и компактность

Многие работы по распознаванию образов опираются на основополагающую гипотезу, названную ее авторами гипотезой о компактности образов. Принципы самообучения распознаванию образов, рассматриваемые далее, вытекают из этой гипотезы. Они таковы: если пространство, определяемое заданными признаками, не обладает свойством компактности, то самообучение не возможно, и наоборот, если в выбранном пространстве гипотеза компактности выполняется, то самообучение становится возможным. Другое дело, что достигнуть нужного разбиения не всегда удается, так как оно зависит от выбранной метрики, и если эта метрика выбрана неудачно, то таким же будет и разбиение.

Назовем некоторое случайно выбранное пространство абстрактным, а отображение любого объекта на это пространство – абстрактным изображением. В любом абстрактном пространстве могут существовать компактные множества точек. Поэтому множество объектов, которым в выбранном абстрактном пространстве соответствует компактное множество точек, можно назвать абстрактным образом данного пространства. Из интуитивных представлений, связанных с понятием образа, следует, что изображения, лежащие достаточно близко к тому или иному изображению из некоторого образа, должны принадлежать тому же образу. Это обстоятельство должно быть учтено при определении понятия абстрактного образа.

Абстрактным образом V_i будем называть любое множество точек X_i, $(i = 1, 2, \ldots)$, для каждой из которых выполняется неравенство

$$\min_{X_k \in v_j} d(X_i, X_k) < \min_{X_\nu \notin X_j} d(X_i, X_\nu), \qquad (1)$$

где $d(X_i, X_k)$ и $d(X_i, X_\nu)$ – некоторая мера близости между абстрактным изображением X_i и изображениями X_k и X_ν соответственно. Так как точкам X_i соответствуют некоторые объекты, будем считать, что эти объекты принадлежат одному абстрактному образу.

Множество элементов $\overline{X} \in X$ в своей совокупности назовем обучающей последовательностью.

Приведенное неравенство является естественным критерием разбиения, требующим, чтобы мера близости между любой точкой из абстрактного образа и ближайшей к ней точкой из того же образа была меньше меры близости между этой же точкой и ближайшей точкой из другого абстрактного образа. Естественный критерий дает формальное определение не только абстрактного образа, но и ранее определенного интуитивного понятия компактного множества точек в абстрактном пространстве.

Многие из известных алгоритмов самообучения способны выделять без подсказок только абстрактные образы в заданных пространствах, т.е. компактные множества заданного пространства. Различия между ними состоят в формализации понятия компактности. Если реальным образам, т.е. не зависящим от выбора пространства, будут соответствовать компактные множества в выбранном пространстве, то понятия реального и абстрактного образов совпадут.

Процесс выработки способности к выделению в заданном пространстве заранее неизвестного количества абстрактных образов только по данным смешанной обучающей выборки называется *самообучением распознаванию образов*. Как правило, процесс самообучения упрощается заданием числа образов.

Несмотря на то что процесс самообучения может привести только к выделению абстрактных образов, это не снижает (а иногда и повышает) ценность таких алгоритмов, так как реальные образы совпадают с абстрактными. Кроме того, очень часто образы в заданном пространстве никем не определены, а задача состоит в том, чтобы в этом пространстве выделить группы похожих объектов. Примерами могут быть многие задачи технической и медицинской диагностики и социологии, в которых учитель не знает четкой классификации анализируемых объектов, а знает только пространство (совокупность признаков) желательного проведения классификации.

Абстрактные образы не всегда совпадают с реальными образами, называемые образами реального мира, существующими вне нас и независимо от нас. Поэтому реальными образами лучше называть образы, указанные учителем.

Две группы алгоритмов

Любая процедура обучения (или самообучения) распознаванию образов состоит в установлении на множестве X такого отношения, которое порождает разбиение этого множества на подмножества, соответствующие образам. Найденное отношение характеризует некоторые особые свойства изображений, по которым они группируются в образы, т.е. свойства принадлежности изображений. На основе этого строится некоторая решающая функция, определяющая принадлежность любого изображения из X.

В зависимости от того, где определяется решающее правило, алгоритмы обучения и самообучения можно разделить на две группы:

- алгоритмы определения решающего правила на множестве аргумента, т.е. на множестве элементов из X;
- алгоритмы определения решающего правила на множестве функций.

К первой группе относятся все алгоритмы построения в пространстве X разделяющих поверхностей, а ко второй – алгоритмы, в основе которых лежит восстановление некоторых функций, обладающих особыми свойствами, указывающими принадлежность объектов к тому или иному компактному множеству. И в тех и в других алгоритмах осуществляется поиск некоторого свойства принадлежности, определяющего отношение на множестве X. Так, например, если в пространстве построена некоторая поверхность, то свойство принадлежности определяется взаимным расположением изображений и этой поверхности. Если же восстановлены условные распределения вероятностей, то свойство принадлежности определяется в зависимости от соотношения этих функций.

Зададим на множестве аргумента такую функцию, что каждому элементу X ставится в соответствие только один элемент другого множества P. Если в качестве задаваемой функции используется функция плотности вероятности $P(X)$, определенная каким-либо способом, то множество P представляет собой совокупность всех точек этой функции.

Особенность рассматриваемого метода состоит в том, что исходное множество X (множество аргумента) отображается на множестве значений функции и каждый элемент нового множества P сразу же приобретает свойство принадлежности, состоящее в том, что каждая точка множества P наделяется некоторой силой притяжения, направленной к единственному центру притяжения. И если каждая точка множества значений функции P будет представлена самой себе, то она, перемещаясь под воздействием силы притяжения, совместится со своим единственным центром притяжения. Функция P должна быть выбрана так, чтобы каждому абстрактному образу соответствовал единственный центр притяжения, а все точки каждого абстрактного образа наделялись бы силой притяжения, направленной к своему центру.

Очевидно, что функцию с требуемыми свойствами построить невозможно, но оказывается, что она объективно существует. Именно такой функцией есть функция плотности вероятности смеси P (ФПВ смеси).

Предположим, что известны все статистические свойства множества X, в частности $P(V_i)$ и $P(X)/V_i$., где V_i – образы, $i = 1, 2, \ldots$ $\ldots, m$. Рассматривается множество X как множество точек в пространстве заданных признаков. В этом случае задача распознавания образов не вызывает никаких принципиальных затруднений. Предположим далее, что задача распознавания решается методом максимального правдоподобия по алгоритму

$$X \in V_{k,} \text{ если } P(X / V_k) = \max_i P(X / V_i), \quad (2)$$

где $(P(X)/V_i)$ – нормальные распределения с одинаковыми дисперсиями. В этом идеализированном случае равенство (2) для образа V_k соблюдается в некоторой окрестности M_k (M_k – математическое ожидание распределения $(P(X)/V_k)$, а в некоторой окрестности точки $X = M_k$ функция $P(X)$ достигает локального экстремума-максимума, т.е. в окрестности, ограниченной «оврагом» вокруг указанного экстремума на $P(X)$, выносится решение в пользу V_k. Если же $(P(X)/V_i)$ отличаются от нормальных распределений и не совпадают по форме, а образы V_i неравновероятны, то границы, определяющие решения в пользу различных классов, не совпадают точно с линиями «оврагов» на $P(X)$, но все же проходят в окрестностях этих «оврагов». Таким образом, если бы удалось каким-либо способом восстановить линии «оврагов» на $P(X)$, то можно было бы построить процедуру распознавания, которая мало отличалась бы по своим результатам от процедуры, организованной по алгоритму максимального правдоподобия при известных $P(X/V_i)$.

Но функция плотности вероятности $P(X)$ (ФПВ смеси) может быть восстановлена и по смешанной представительной обучающей последовательности, в которой отсутствуют указания о принадлежности изображений различным образам.

В работе [4] указывается, что решение задачи самообучения может быть сведено к восстановлению ФПВ смеси и определению по ней «центров», а затем и границ образов; однако автор отмечает, что задача построения решающей функции, определяемой границей, проходящей по «дну оврагов» ФПВ смеси, от этого не упростилась. Действительно, если использовать ФПВ смеси только для того, чтобы спроектировать ее «овраги» на пространство и полученные проекции принять за разделяющие границы, то это не упрощает задачу самообучения в целом. Но оказывается, эту границу и не нужно искать. Если считать, что ФПВ смеси уже восстановлены, то задачу самообучения можно считать решенной, так как требуемое отображение исходного множества аргументов X на множество функции плотности вероятности уже совершено и система, которая может вычислять значения ФПВ в любой точке, уже приняла требуемую организацию.

Предположим, что ФПВ смеси восстановлена. Отобразим множество аргументов X на множество значений ФПВ. Неравномерное распределение элементов X отобразится в многомодальную структуру ФПВ. Каждая область «сгущения» элементов X отобразится в область значений ФПВ, тяготеющих к одной моде. Именно на этом свойстве множества аргументов и множества значений ФПВ основан предлагаемый подход к решению задачи самообучения.

Метод смешанных распределений

Пусть на вход обученной системы, т.е. системы, в памяти которой хранятся значения ФПВ для любой точки X, предъявляется объект с неизвестной заранее классификацией. Найдем отображение этого элемента на множестве функций $P(X)$ и организуем поиск локального экстремума-максимума ФПВ в окрестности этой точки. Все отображения, лежащие по одну сторону от «дна оврага» на ФПВ, будут «скатываться» к одному экстремуму (моде), т.е. будут тяготеть к одной моде ФПВ. Если считать реакцией системы указание номера моды, к которой тяготеют элементы X, то система будет отвечать одинаковой реакцией на все множество элементов и соответствующее подмножество объектов, тяготеющих к одной и той же моде.

В соответствии с этим метод смешанных распределений формулируется следующим образом: в процессе самообучения восстанавливается только оценка ФПВ смеси; в процессе распознавания для каждого объекта вычисляется значение ФПВ смеси, и в окрестности этой точки на ФПВ организуется поиск локального экстремума-максимума. Результатом распознавания есть указание координат, соответствующих точке найденного экстремума; далее задача состоит в том, чтобы интерпретировать полученные результаты распознавания, т.е. каждому абстрактному образу присвоить удобное название.

Сущность такого подхода к решению задачи самообучения состоит в том, что выборке объектов с неизвестной заранее классификацией сначала представляется возможность проявить особенности своего распределения и на основании этих особенностей выработать свойство принадлежности элементов X.

Методом смешанных распределений можно проводить автоматическую классификацию в соответствии с объективно существующим распределением абстрактных образов в пространстве, определяемом заданной метрикой. Любое существенное изменение в структуре абстрактных образов находит свое отображение на множестве значений ФПВ смеси, вызывая соответствующие изменения во взаимном расположении экстремумов, их количестве, конфигурации «оврагов», разделяющих экстремумы-максимумы, и т.д. Метод смешанных распределений не требует существенных ограничений на вид условных распределений $P(X)/V_i$, кроме одномодальности этих функций.

Методы оценки ФПВ смеси

Успех решения задачи самообучения по методу смешанных распределений зависит, прежде всего, от качества оценки ФПВ в любой точке заданного пространства. Поэтому рассмотрим этот вопрос более подробно.

Так как функциональный вид распределения смеси заведомо неизвестен и может быть достаточно сложным, невозможно принять какие-либо обоснованные предположения относительно функционального вида ФПВ смеси, которые могли бы если не сразу, то во всяком случае при малом числе конкурирующих гипотез подтвердиться с помощью самой выборки наблюдений. Поэтому в [6] для оценки ФПВ смеси применяются непараметрические методы, т.е. не зависящие от функционального вида распределения смеси.

Рассмотрим непараметрические методы оценки ФПВ с использованием различных функций вкладов. Суть этих методов состоит в том, что с каждой точкой X_i из смешанной обучающей последовательности связывается некоторая функция вклада $P_i(X_i, X)$, где X_i и X – точки в пространстве, X_i – фиксированная точка из обучающей последовательности, а X – любая другая точка пространства, называемая текущей точкой.

Оценка ФПВ в точке X определяется в виде суммы

$$\hat{P}_N(X) = \frac{1}{N}\sum_{i=1}^{N} P_i(X_i, X), \qquad (3)$$

где N – число точек в обучающей последовательности. Функция вклада обычно выбирается как некоторая функция расстояния между X_i и X.

Методы вкладов очень близки по своему содержанию к методу потенциальных функций [3].

Алгоритмы метода вкладов обеспечивают некоторые специальные свойства оценок ФПВ, необходимые для применения метода смешанных распределений при ограниченной обучающей последовательности.

В работе [6] подробно рассматриваются два метода, названные методами нормальных вкладов и ближайшего наблюдения. В методе нормальных вкладов функция вклада имеет вид

$$P_i(X_i, X) = \frac{1}{\sigma_i}\exp\left(-\frac{d(X_i, X)}{2\sigma_i^2}\right), \qquad (4)$$

где σ_i – размах вклада; $d(X_i, X)$ – расстояние между точками X_i и X. X_i – точка обучающей последовательности; X – любая точка пространства. В этой же работе исследуются асимптотические свойства оценок [3] и показывается, что по мере роста N эти оценки сходятся к действительным значениям ФПВ смеси. Оценки ФПВ смеси, полученные по методу нормальных вкладов на конечных выборках, как и по другому непараметрическому методу, основанному на использовании функций вкладов, как правило, смещены и несостоятельны. Поэтому важно указать ограничения на параметр σ_i нормального вклада для конечных N, чтобы получаемая при этом оценка ФПВ смеси обладала нужными свойствами. Показано, что для того чтобы получить удовлетворительные оценки ФПВ смеси, позволяющие применить для автоматической классификации метод нормальных вкладов необходимо каждой точке обучающей выборки поставить в соответствие свой параметр вклада σ_i, где

$$\sigma_i \geq \min_{k\neq 1} d(X_i, X_k); \quad k = 1, 2, ..., N. \qquad (5)$$

После этого по формуле (5) можно определить значение оценки ФПВ смеси в каждой точке заданного пространства.

Заключение. Оценка ФПВ, так же как и сама функция $P(X)$, должна иметь единственный экстремум-максимум над каждым абстрактным образом; при этом координаты этих экстремумов не обязательно должны совпадать. Оценка ФПВ смеси над каждым абстрактным образом должна быть непрерывной, дифференцируемой и не иметь дополнительных (ложных) экстремумов. Кроме того, «овраги» оценки ФПВ смеси должны проходить в окрестности «оврагов» на функции $P(X)$. Если оценка ФПВ смеси сохраняет эти свойства функции $P(X)$, то ее называют равноэкстремальной. При сохранении свойства равноэкстремальности оценок ФПВ смеси для самообучения применим метод смешанных распределений.

Окончание на стр. 46

Для сохранения свойства равноэкстремальности оценки ФПВ необходимо наложить определенные ограничения на вид функции вклада и на свойства обучающей последовательности.

Функция вклада должна иметь «гладкий» экстремум, т.е первая производная в области экстремума функции вклада должна плавно проходить через нуль. Для этого надо с каждой точкой обучающей последовательности связать функцию вклада (5). Не соблюдение этого условия может быть причиной образования дополнительных экстремумов-максимумов, что приведет к выделению нескольких групп векторов в одном абстрактном образе. Для ряда практических задач это оказывается несущественным, так как такие области образуют область одного класса.

Более подробный перечень требований к восстанавливаемым ФПВ смеси приведен в [6].

1. *Розенблатт Ф.* Принципы нейродинамики (перцептрон и теория механизмов мозга). – М.: Мир, 1965. – 480 с.
2. *Ивахненко А.Г.* Самонастраивающиеся системы с положительными обратными связями. – Киев: Изд-во АН УССР, 1963. – 378 с.
3. *Айзерман М.А.*, *Браверманн Э.М.*, *Розоноэр Л.И.* Метод потенциальных функций в теории обучения машин. – М.: Наука, 1970. – 384 с.
4. *Цыпкин Я.З.* Адаптация обучения в автоматических системах. – М.: Наука, 1968. – 400 с.
5. *Шлезингер М.И.* Синтез линейного решающего правила для одного класса задач распознаванию образов // Изв. АН СССР. Сер. Техн. кибернетика. – 1972. – № 5. – С. 129–135.
6. Васильев В.И. Распознающие системы: Справочник. – Киев: Наук. думка, 1983. – 419 с.

Поступила .24.01.2011
Тел. для справок: (044) 526-4187, 234-7721 (Киев)

●