DOI: 10.25728/mlsd.2023.1398

КОРРЕЛЯЦИОННЫЙ МЕТОД СВЕРТКИ ИЗОБРАЖЕНИЙ

Зенков В.В.
Институт проблем управления им. В.А. Трапезникова РАН, Москва, Россия
zenkov-v@yandex.ru

Аннотация. Предложен метод свертки изображений, основанный на корреляции элементов изображений обучающей выборки с учителем с аппроксимируемой функцией, используемой для классификации. Приведен пример свертки изображений, используемых для классификации рукописных цифр путем оценки апостериорных вероятностей десяти классов изображений с помощью дискриминантной функции Андерсона.

Ключевые слова: машинное обучение, классификация изображений, свертка изображений, обучающая выборка с учителем, дискриминантная функция Андерсона, апостериорная вероятность класса, коэффициент корреляции.

Введение

При классификации изображений нейросетевым методом используется операция свертки изображений для упрощения задачи и сокращения времени работы программы. Для решения задач классификации изображений простым и естественным методом, использующим аппроксимацию дискриминантной функции Андерсона (ДФА) также приходится выполнять свертку изображений.

В основе применения ДФА для решения задач классификации лежит простая модель мыслительной деятельности человека: не слоеный пирог нейронов, а совокупность близких к решаемой задаче случаев из практики отдельного человека или коллектива людей, называемых в машинном обучении обучающей выборкой с учителем.

ДФА, по знаку которой объект относится в один из двух классов, непосредственно определяет и апостериорные вероятности (АпоВ) принадлежности объекта к классам, что дает возможность решать задачу классификации по АпоВ и по различным субъективным критериям, в том числе и в виде многокритериальной задачи. Формула вычисления по ДФА в точке пространства признаков объектов классификации вероятностей классов, которым может принадлежать объект классификации, не сложнее формулы вычислений градусов Фаренгейта по градусам Цельсия.

Если классов больше двух, используется метод один класс против остальных, принимаемых за второй класс. Он позволяет последовательно найти АпоВ всех классов для классифицируемого объекта и по ним отнести объект в нужный класс по выбранному критерию или отказаться от решения, объяснив принимаемое действие.

Для аппроксимации ДФА в непараметрическом методе не нужно задаваться видом аппроксимирующей функции, достаточно выбора линейной зависимости в окрестности заданной точки в пространстве вещественных признаков, характеризующих объект классификации. Таким способом построить касательную к ДФА в точке. И совсем просто получать аппроксимацию ДФА в точке способом Надарая-Ватсона в виде средневзвешенной оценки ДФА в точке.

Для непараметрического метода аппроксимации ДФА обучающая выборка нужна не только в период обучения. Она нужна и в дальнейшем, но в свернутом виде. Кроме того, нужно определяться с мерой расстояния между точками в признаковом пространстве, с видом весовой функции и, обычно, с единственным ее параметром настройки в отличие от гигантского количества настраиваемых коэффициентов связи, выбора подходящих функций активации нейронов и градиентного метода поиска экстремума целевой функции сложного вида в пространстве большой размерности, что свойственно нейросетевым методам.

Математический аппарат, используемый для работы с ДФА, – взвешенный метод наименьших квадратов. Для подбора единственного параметра весовой функции можно ограничиться простым перебором значений из заданного диапазона.

При решении примеров задач классификации [1- 3] в качестве весовой функции между заданной точкой и точкой обучающей выборки использовалась экспонента от эвклидова расстояния между точками, умноженного на отрицательный коэффициент настройки. С увеличением количества точек обучающей выборки, называемого нейросетевиками «глубоким» обучением, этот коэффициент настройки увеличивается по модулю, приближая процесс оценивания АпоВ одного из двух классов по сути к многократному подбрасыванию монеты, когда роль одной монеты играет «множество близких» точек выборки к заданной точке.

Решаемые нами ранее задачи классификации не имели дела с изображениями. Пространства признаков имели небольшие размерности и количества точек в обучающих выборках были сравнительно небольшими по сравнению с «глубоким» обучением. Рекордным по минимальному количеству точек являлась медицинская задача классификации на 3 класса с обучающей выборкой из 36 точек [3].

Особенностью задач классификации изображений являются сравнительно гигантские размерности пространств признаков. В качестве признаков выступают пиксели или группы пикселей изображений. Так, в приводимом далее примере распознавания рукописных цифр MNIST (Modified National Institute of Standards and Technology), взятого из интернета [4], обучающая выборка использует матрицу изображения размером 28x28, т.е. 784 признаков –пикселей с оттенками серого цвета из диапазона 0-255. Для обучения классификатора с 10 классами, с тестовой выборкой размером 10000 объектов и с обучающей выборкой в 60000 объектов программе, написанной на Питоне и не использующей ни распараллеливания, ни графического процессора, потребуется на небольшом компьютере много часов работы.

Для свертки изображений при решении подобных задач предложен и испытан корреляционный метод свертки для последующей аппроксимации ДФА и прочих регрессионных зависимостей. Матрица изображения преобразуется в строку признаков-пикселей. По обучающей выборке находятся корреляции признаков с искомой величиной и корреляции между признаками. Затем отбрасываются слабо коррелированные с искомой величиной признаки. Кроме того, из двух сильно коррелированный между собой признаков отбрасывается признак, менее коррелированный с искомой величиной. Два задаваемых для этого предельных значений коэффициентов корреляции выбираются из условия получения приемлемого количества используемых далее признаков, позволяющих решать задачу обучения и классификации в приемлемое время с приемлемой точностью. Количество параметров настройки, таким образом, увеличивается до трех.

В случае с количеством классов больше двух (в нашем примере их 10), когда решается несколько задач по принципу один класс против всех остальных, принимаемых за второй класс, можно получать свертку для каждого класса, принимаемого за первый.

Если имеет место несбалансированность обучающей выборки – существенное различие в количествах точек разных классов, то можно использовать метод [5] с выбором стоимостей ошибок классификации для обеспечения субъективно приемлемых результатов классификации.

1. Дискриминантная функция Андерсона и апостериорные вероятности классов

ДФА в случае с двумя классами мы определяем, как разность двух функций средних стоимостей потерь от ошибок при отнесении точки x пространства признаков классов (объекта классификации) либо к одному, либо к другому классу. Называем эту функцию дискриминантной функцией Андерсона в честь Теодора Уилбура Андерсона, автора простого способа решения байесовой задачи классификации [6]. ДФА – это разность двух функций средних потерь от ошибок классификации:

$$f_{01}(x, C) = G_0(x, C) - G_1(x, C) = \mathrm{M}_{k|x}(C_{ok} - C_{1k}), \qquad (1)$$

где $G_0(x,C)=C_{01}\left(1\text{-}p(0|x)\right)$, $G_1(x,C)=C_{10}p(0|x)$–средние потери в точке x, если точку относить к классу 0 или, соответственно, к классу 1,

C_{ij} –стоимость ошибки, когда точка из класса j ошибочно относится в класс i, $C_{01}>0, C_{10}>0$, $C_{ii}=0$,

$p(k|x)$ – АпоВ класса k в точке x, $p(0|x) + p(1|x) = 1$,

$p(k|x) = \dfrac{P_k p(x|k)}{p(x)}, p(x) = P_0 p(x|0) + P_1 p(x|1)$

P_k, –априорные вероятности классов,

$p(x|k)$ – условные распределения признаков классов,

k –номер класса 0 или 1,

$\mathrm{M}_{k/x}(\cdot)$ –математическое ожидание по k в точке x.

Если $f_{01}(x, C) < 0$, то точку x следует отнести в класс 0, иначе – в класс 1.

1.1. Свойства

Перечислим свойства ДФА, опустив доказательства из-за их очевидности:

1.ДФА по определению есть функция регрессии (1). Для преобразования обучающей выборки задачи классификации с учителем в выборку регрессионного анализа нужно заменить в выборке номер класса k=0 на минус C_{10}, а номер класса k=1 на +C_{01}, учитывая C_{11}=C_{22}=0 в (1).

2.ДФА есть ограниченная функция регрессии, $-C_{10} < f_{01}(x, C) < C_{01}$.

3.Для АпоВ нулевого класса и ДФА, определенной для заданной матрицы стоимостей ошибок классификации C, имеет место тождество, вытекающее из (1),

$$p(0|x) \equiv (C_{01} - f_{01}(x, C))/(C_{01} + C_{10}). \qquad (2)$$

4.Если задавать стоимости ошибок при условии $(C_{01} + C_{10}) = 1$, что не приводит к потере общности, то тождество (2) не только предельно упростится,

$$p(0|x) \equiv p^* - f_{01}(x, p^*), \qquad (3)$$

но и стоимости ошибок обретут удобный для интерпретации смысл

$$C_{01} = p^*,\ C_{10} = 1 - p^*, \qquad (4)$$

где p* - это еще и АпоВ первого класса на границе между классами в пространстве признаков, где $f_{01}(x, p^*) = 0$, если граница существует при заданной p^* и имеющейся обучающей выборке.

5.Результат вычисления АпоВ по (3) **не зависит от выбора значения p^*** и существования границы между классами. АпоВ класса является объективной величиной, в то время как параметр p^* - величина субъективная, произвольно задаваемая для определения ДФА.

6.Отбор признаков или функций от признаков для аппроксимации ДФА может производиться по коэффициентам корреляции с оценками ДФА, $C_{0k} - C_{1k}$, с исключением одного из пары отобранных, но слишком коррелированных между собой признаков.

7.Поскольку ДФА является функцией регрессии от x, то и АпоВ класса в соответствии с (3) также есть функция регрессии от x. Поэтому АпоВ класса можно строить совершенно аналогично ДФА, обозначая в обучающей выборке в соответствии с (3) один класс нулем, а другой – единицей.

2. Классификация рукописных цифр

В отличие от представления объекта классификации в виде набора действительных чисел – признаков графическое изображение объекта представляется матрицей из нескольких строк и столбцов. Для использования метода классификации, отлаженного на неграфических объектах, для классификации графических нужно матрицу графического объекта развернуть в строку признаков. Значениями признаков графического объекта будут цветовые и оттеночные коды элементов исходной матрицы, используемых в качестве признаков объекта классификации. Для уменьшения размерности пространства признаков классов необходимо произвести свертку - отбор информативных признаков.

В качестве меры информативности признака в регрессионной задаче предлагается использовать коэффициент корреляции признака с ДФА при ограничении на коррелированность признаков между собой. Эти два параметра вместе с параметром весовой функции являются тремя параметрами настройки, определяемые в процессе обучения.

Предложенный способ опробован на популярной задаче классификации рукописных цифр MNIST. Данные получены известными средствами Питона [4]. Обучающая выборка содержит 60000 изображений рукописных цифр в среднем по 6000 изображений. Матрица изображений каждой цифры 28x28 преобразуется в строку с 784 действительными значениями оттенков серого в диапазоне от 0 до 255. Набор тестовых изображений составляет 10000 строк в среднем по тысяче изображений на цифру.

Имея компьютер с 64 разрядным процессором Intel(R) Core(TM) i5-9400 CPU @ 2.90GHz 2.90 GHz , с оперативной памятью 8 GB с операционной системой Windows 10 Pro экспериментально остановились на выборе варианта свертки с предельно минимальным модулем корреляции признака с ДФА 0.37 и с предельно максимальным модулем корреляции признаков между собой 0.25. Эти два параметра передавались программе, написанной на Питоне. Кроме того, задавался диапазон, в котором должно находиться количество отобранных признаков - от 30 до 50 из 784-х. Если количество признаков превышало 50, программа уменьшала верхний предел корреляции признака с ДФА, например, 0.37 с шагом 0.05 и т.д. Если количество становилось меньше 30, то программа увеличивала полученный предел на 0.02. Равными размеры шагов уменьшения и увеличения брать нельзя во избежание зацикливания программы. За несколько итераций количество признаков загонялось в коридор 30-50. Предельная корреляция между признаками 0.25 не изменялась.

В Таблице 1 представлены результаты работы при лучшем значении коэффициента весовой функции -10, минимизирующем среднюю ошибку классификации по тестовой выборке. Он найден перебором по нескольким значениям коэффициента весовой функции. Зависимость средней ошибки от этого коэффициента представлена в Таблице 2. В Таблице 3 представлены отобранные признаки для разных классов в графическом представлении 28x28.

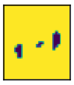
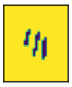
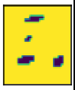
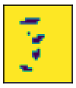
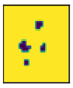
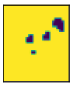
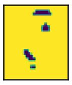
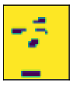
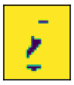
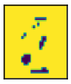
Таблица 1.

Цифра	**Кол. точек**	**Кол. ошибок**	**Процент ошибок**	**Кол. признаков**	**Корр. с ДФА**
0	980	39	4.0	32	0.37
1	1135	15	1.3	35	0.29
2	1032	202	**19.6**	46	0.24
3	1010	123	12.2	39	0.22
4	982	106	10.8	30	0.22
5	892	138	15.5	36	0.17
6	938	96	10.0	34	0.27
7	1028	140	13.6	49	0.24
8	974	190	**19.5**	33	0.17
9	1009	130	12.9	42	0.19

Таблица 2.

Коэффициент	-3.0	-5.0	-7.0	-10.0	-15.0	-20.0
Ошибка в %	21.3	15.0	12.4	11.8	12.2	12.6

Таблица 3.

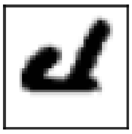
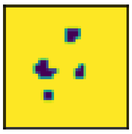
Цифра	0	1	2	3	4	5	6	7	8	9
Свертка										

Из Таблицы1 видно, что величины ошибок в классификации рукописных цифр очень различны и лежат в пределах от 1.3 % для цифры 1 и до 19.6 % для цифры 2. Возможно, что для выравнивания результатов следует предельные корреляции с ДФА для некоторых проблемных изображений задавать индивидуально. Например, по Таблице 1 такими проблемными изображениями нужно считать изображения цифр 2 и 8. В целом, учитывая небольшое количество признаков после свертки из исходного количества 784 признаков, можно результат решения задачи считать удовлетворительным.

Время расчета ошибок, представленных в Таблице 1, с преобразованием исходной обучающей выборки в выборку из результатов свертки по 10 классам и с расчетом ошибок 10000 изображений тестовой на вышеописанном компьютере составило 32 часа. Для заполнения Таблицы 2 задача запускалась 6 раз.

При демонстрации работы обученной программы предлагалось ввести номер изображения из тестовой выборки от 0 до 9999, например, 8000. Появлялось в ответ соответствующее изображение рукописной цифры вместе с изображением свертки, использованной при классификации для этой цифры в режиме один против всех:

Введите номер изображения из диапазона 0-9999: 8000

Изображение и фигура свертки для истинного класса 4.

Затем программа вычисляла АпоВ всех 10 классов в таком же режиме со своими свертками, Таблица 3, по обучающей выборке и через 10 секунд выдавала принятое решение верное или ошибочное и все АпоВ классов:

Вероятности по классам один против всех
['0.011', '0.000', '0.000', '0.000', '0.881', '0.000', '0.000', '0.100', '0.000', '0.000']
Фигура отнесена в класс 4.
Введите номер изображения из диапазона 0-9999:

3. Заключение

1. Аппроксимация дискриминантной функции Андерсона – это простой и легкий способ оценки апостериорных вероятностей классов по обучающей выборке с учителем. Использование апостериорных вероятностей классов позволяет решать задачу классификации в стохастической постановке в два этапа. На первом, объективном, этапе для объекта классификации по его признакам находятся апостериорные вероятности классов, к которым он может принадлежать. На втором, субъективном, этапе с учетом заданных стоимостей ошибок классификации и прочих условий выполняется собственно классификация объекта.
2. При классификации графических объектов из-за большого количества признаков необходимо производить свертку изображений – отбор заданного количества признаков, упрощающих и ускоряющих решение задачи классификации без существенного ухудшения качества решения.
3. Предложен простой метод свертки изображений, использующий корреляционные связи признаков с регрессионными зависимостями типа дискриминантной функции Андерсона, используемыми для классификации изображений прямо или посредством оценивания апостериорных вероятностей классов, которым может принадлежать изображаемый объект.
4. Приведен пример решения задачи классификации изображений рукописных цифр. Свертка изображений выполнялась корреляционным методом. Классификация выполнялась путем оценивания апостериорных вероятностей классов по дискриминантным функциям Андерсона, получаемым непараметрическим методом.
5. Сравнительно невысокое качество распознавания некоторых цифр вполне объяснимо поставленной задачей – продемонстрировать метод свертки на малом количестве выделенных в результате свертки признаков (30-50 из 784) из-за ограничения времени обучения на имеющемся компьютере. При наличии более совершенного компьютера с более совершенной программой, использующей распараллеливание, как того непременно требует «глубокое обучение», можно получить более качественный результат.

Литература

1. *Зенков В.В.* Оценка апостериорной вероятности класса по серии дискриминантных функций Андерсона // АиТ. 2019, № 3. С. 68–82.
2. *Зенков В.В.* Применение аппроксимации дискриминантной функции Андерсона и метода опорных векторов для решения некоторых задач классификации // АиТ. 2020, № 1. С. 147–160.
3. *Зенков В.В. и др.* Оценка апостериорных вероятностей классов в задаче дифференциальной диагностики заболеваний слизистой оболочки полости рта // “Управление развитием крупномасштабных систем” (MLSD’2020): труды Тринадцатой междунар. конф: – М.: ИПУ РАН, 2020. – С.1793-1801.
4. Создаем программу для распознавания рукописных цифр с tensorflow и tkinter https://pythonru.com/primery/raspoznavanie-rukopisnyh-cifr-na-p-ython-gui
5. *Зенков В.В.* Применение аппроксимации дискриминантной функции Андерсона и апостериорных вероятностей классов при несбалансированной обучающей выборке в машинном обучении // “Управление развитием крупномасштабных систем” (MLSD’2022): труды Пятнадцатой междунар. конф: – М.: ИПУ РАН, 2022. – С.1209-1215.
6. *Anderson T. W.* An Introduction to Multivariate Statistical Analysis. –Third Edition. Hoboken, NJ: John Wiley & Sons, 2003. –742 pp.