УДК 004.048

ОЦЕНКА ВЕРОЯТНОСТЕЙ ПРИНАДЛЕЖНОСТИ ТОЧКИ КЛАССАМ ПО АППРОКСИМАЦИЯМ ДИСКРИМИНАНТНЫХ ФУНКЦИЙ АНДЕРСОНА

В.В. Зенков
Институт проблем управления им. В.А. Трапезникова РАН
Россия, 117997, Москва, Профсоюзная ул., 65
E-mail: zenkov-v@yandex.ru

Ключевые слова: машинное обучение, классификация, оценка апостериорной вероятности класса, аппроксимация дискриминантной функции, дискриминантная функция Андерсона, калибратор Платта.

Аннотация: Предложены два способа оценки апостериорных вероятностей классов в заданной точке пространства признаков объектов по аппроксимациям дискриминантных функций Андерсона от признаков в окрестности нулевых значений функций. В точках нулевых значений дискриминантной функции Андерсона апостериорные вероятности классов зависят только от стоимостей ошибок классификации, для которых определена эта функция. Аппроксимации строятся по обучающей выборке с учителем. Апостериорные вероятности классов используются для объяснения принимаемых решений при классификации объектов, позволяя решать задачу классификации для различных критериев и различных стоимостей ошибок классификации. Способы не нуждаются в надстройках типа калибратора Платта для метода опорных векторов.

1. Введение

В середине прошлого века для решения задач классификации усилия исследователей были сосредоточены на поиске методов, обеспечивающих лучшие средние показатели классификации. И только к концу столетия стали появляться работы, авторы которых стали обращать внимание на объяснение принимаемых решений при отнесении конкретного объекта в тот или иной класс. Наиболее убедительными доводами для этого служат апостериорные вероятности (АпоВ) возможной принадлежности объекта к классам, рассчитанные по совокупности показателей (признаков), присущих объекту. Так появились надстройки типа калибратора Платта [1] для пересчета отступов от границ классов в случае использования для классификации дискриминантных функций (ДФ). Таким образом, после решения оптимальной в среднем задачи классификации приходится решать дополнительную задачу для получения оценок АпоВ, необходимых для надежных объяснений принимаемых решений при отнесении объектов в тот или иной класс.

Наряду с этим есть возможность по обучающей выборке с учителем напрямую находить оценки АпоВ классов в заданных точках пространства признаков без решения дополнительных задач, и это не связано с восстановлением условных распределений признаков классов и использованием формулы Байеса. Достаточно лишь использовать

ДФ Андерсона (ДФА) [2] и ее аппроксимацию (АДФ) [3] в окрестности нулевых значений, где АпоВ классов определяются только стоимостями ошибок классификации, которые входят в определение ДФА. Математический аппарат, используемый для решения задачи – взвешенный метод наименьших квадратов. Размерность обращаемой матрицы при решении задачи не зависит от объема обучающей выборки, она определяется количеством используемых функций от признаков, входящих в АДФ.

2. Дискриминантная функция Андерсона и ее свойства

Определение ДФ Андерсона получим из способа оптимального решения байесовой задачи классификации для случая нескольких (более двух) классов, описанного Т. Андерсоном [4]. Для каждого класса в заданной точке пространства признаков по апостериорным вероятностям классов подсчитываются средние потери от отнесения точки в класс. Точка относится в тот класс, для которого средние потери в точке меньше, чем для других классов.

Определение. *ДФ в виде разности средних потерь двух классов будем называть ДФ Андерсона.*

Если ДФА в заданной точке меньше нуля, то точку относим в один класс, иначе – в другой. Далее нам будет достаточно ограничиться случаем с двумя классами. Если классов больше двух, то для оценки АпоВ классов потребуется решать несколько задач с двумя классами по принципу один класс против всех остальных классов. Возможно при этом потребуется использовать более сложный вид функций, используемых для аппроксимации ДФА.

Средние потери $G_1(x)$ и $G_2(x)$ от отнесения точки x в первый и второй классы

$$G_1(x) = C_{12} p(2\,|\,x), \quad G_2(x) = C_{21} p(1\,|\,x), \quad p(1\,|\,x) + p(2\,|\,x) \equiv 1,$$

где C_{12} – стоимость потери, если точка второго класса ошибочно относится в первый класс; C_{21} – стоимость потери, если точка первого класса ошибочно относится во второй класс, $C_{11} = C_{22} = 0$; $p(1|x)$ и $p(2|x)$ – АпоВ первого и второго классов, их сумма равна единице. ДФА по определению:

(1) $$\mathrm{f}_{12}(x) = G_1(x) - G_2(x) = \mathrm{M}_{k|x}\left(C_{1k} - C_{2k}\right),$$

где $\mathrm{M}_{k|x}(\cdot)$ обозначает математическое ожидание по номеру класса k в точке x.

2.1. Свойства ДФ Андерсона

Утверждение 1 . ДФА есть ограниченная функция регрессии от признаков.

Доказательство утверждения 1. Факт регрессионной зависимости ДФА очевиден по определению (1), как и ее ограниченность:

$$-C_{21} < f_{12}(x) < C_{12}.$$

***Следствие* 1**. Чтобы аппроксимировать ДФА как функцию регрессии, следует преобразовать обучающую выборку с учителем в задаче классификации в выборку задачи регрессионного анализа, заменив номера классов в выборке следующим образом: первый класс на $-C_{21}$, а второй класс на C_{12} (1).

***Следствие* 2**. Факт регрессионной зависимости ДФА от признаков позволяет, в частности, выполнять и отбор признаков, используемых для решения задачи аппроксимации, по коэффициентам корреляции признаков со столбцом, в котором номера классов заменены стоимостями ошибок классификации. Учитывать при этом необходимо и коэффициенты корреляции признаков между собой.

Утверждение 2. *Для АпоВ первого класса и ДФА, определенной для* C_{12} *и* C_{21}, *име-*

ет место тождество

(2) $$p(1\,|\,x)\ \equiv\ (C_{12} - f_{12}(x))\ /\ (C_{12} + C_{21}).$$

Доказательство утверждения 2. Следует из определения (1) и $p(2\,|\,x)\ \equiv\ 1\ - p(1\,|\,x)$.

***Следствие* 1** . Из (2) видно, что в точках на границе классов, где $f_{12}(x)\ =\ 0$, АпоВ первого класса (обозначим ее через p^*) и стоимости ошибок, для которых определена ДФА, связаны соотношениями

(3) $$p^*\ = C_{12}\ /\ \left(C_{12} + C_{21}\right),\quad 1 - p^*\ = C_{21}\,/\left(C_{12} + C_{21}\right),$$

а в точках, относимых в первый класс, $f_{12}(x)\ <\ 0$, из (2) следует, что $p(1\,|\,x)\ > p^*$.

***Следствие* 2.** Если для ДФА задавать стоимости ошибок при условии $C_{12} + C_{21} = 1$, $C_{12} > 0$, $C_{21} > 0$, что не приводит к потере общности, то тождество (2) не только предельно упростится, но и стоимости ошибок (3) обретут удобный для интерпретации смысл, а именно, $C_{12} = p^*$, $C_{21} =\ 1 - p^*$.

В ДФА для отражения ее зависимости от стоимостей ошибок классификации введем p^* в качестве параметра:

(4) $$p(1\,|\,x)\ \equiv p^*\ - f_{12}(x, p^*).$$

Утверждение 3. *Условия неразличимости классов*: *если в пространстве признаков при заданных стоимостях ошибок* $\min_x\ f_{12}(x, p^*)\ >\ 0$*, то все точки следует относить во второй класс* (*первый класс неразличим*), а *если* $\max_x\ \ f_{12}(x, p^*)\ <\ 0$, *то все точки следует относить в первый класс* (*второй класс неразличим*).

Доказательство утверждения 3. По определению точка относятся в первый класс, если $f_{12}(x, p^*) <\ 0$, что не будет иметь места, если $\min_x\ f_{12}(x, p^*)\ >\ 0$. Аналогично доказывается вторая часть утверждения.

Следствие. Когда имеется не сбалансированная обучающая выборка, когда точек одного класса существенно больше, чем точек другого, то для решения байесовой задачи классификации придется подбирать величину параметра p^*.

Например, нужно уменьшать стоимость ошибки при отнесении точки второго класса в первый класс (точек много), соответственно, увеличивая стоимость ошибки при отнесении точек первого класса, которых мало, во второй класс.

В случае с двумя нормальными условными распределениями признаков классов N(0, 1) и N(m, k^2), m > 0, k > 1 и априорной вероятностью первого класса P_1 в работе [3] получены условия неразличимости первого класса

$$m^2 < 2\ \left(k^2 - 1\right)\ \mathrm{Ln}\left(C_{12}\left(1 - P_1\right)/\left(C_{21}P_1k\right)\right),\quad k\ <\ C_{12}\left(1 - P_1\right)/\left(C_{21}P_1\right)),$$

$$C_{12}\ /\ C_{21} > kP_1\ /\ (1 - P_1).$$

Например, для k=2 и P_1=0,1 задача классификации будет иметь решение, если задавать $p^* < 2/11$, когда m не меньше $6Ln(4{,}5p^*\ /\ (1 - p^*))$.

Если же решать задачу, искусственно уменьшая в выборке количество точек одного из классов, то нельзя будет полученное решение считать решением байесовой задачи классификации при соответствующих оценках АпоВ классов.

3. Оценка апостериорных вероятностей классов

Методы оценивания АпоВ классов в заданной точке основаны на тождестве (2), (4), связывающем АпоВ первого класса с величиной ДФА в этой точке. Поскольку ДФА не известна и имеется лишь обучающая выборка с учителем, в которой номера классов заменены соответствующими разностями стоимостей ошибок, то приходится использовать аппроксимацию ДФА в области нулевых значений, где АпоВ первого класса определяется лишь стоимостями ошибок классификации этой ДФА.

Область нулевых значений не известна, известно лишь, что она находится между двумя уровнями расположения точек выборки. Для аппроксимации ДФА используется прием последовательных приближений к области нулевых значений, использующий взвешенный метод наименьших квадратов [3].

АДФ в виде линейной комбинации заданных функций от признаков $\lambda'\varphi(x)$ с вектором коэффициентов λ находится по обучающей выборке $\{x_n, k_n\}$, $n = 1 \div N$, k_n – номер класса в строке n, $k_n=\{1, 2\}$, $x_n \in R^d$ – вектор действительных значений признаков размерности d в строке n. Решается последовательность задач взвешенным методом наименьших квадратов, минимизирующих по λ_i критерий

$$Q(\lambda_i) = min_{\lambda_i} \sum_{n=1}^{n=N} \{[C_{1k_n} - C_{2k_n} - \lambda_i'\varphi(x_n)]^2 exp(-W_i \,|\, \lambda_{i-1}'\varphi(x_n)\,|)\} \quad (5)$$

где i – номер итерации, i=1 ÷ I. **Н**а первом шаге задача решается без весовой функции, на последующих шагах весовая функция придает больший вес точкам более близким к нулевой области предыдущей АДФ. I – заданное количество итераций, $I < \infty$. W_i – заданный весовой коэффициент, $W_i > 0$, обычно увеличивающийся по итерациям. Это обеспечивает снижение, а потом рост критерия (5) по шагам итераций. Плавность изменения критерия по шагам итераций зависит от шага изменения W_i.

Вид весовой функции - не обязательно экспонента от модуля значений АДФ.

Лучшим значением λ является тот вектор, который дает меньшие потери

$$G = N^{-1}(C_{12}N_2 + C_{21}N_1),$$

где N – количество точек выборки, N_1 – количество точек первого класса, ошибочно отнесенных во второй класс, N_2 – количество точек второго класса, ошибочно отнесенных в первый класс.

3.1. Оценка АпоВ класса по серии АДФ

Для оценки АпоВ классов предварительно по обучающей выборке строится серия АДФ методом (5) по заданным стоимостям ошибок, соответствующим нескольким АпоВ p^*. В процессе построения начальное и конечное значения этих АпоВ могут быть изменены, если окажется, что все точки выборки будут помещаться в один из классов. Если задача классификации имеет больше двух классов, нужно будет строить несколько серий, на единицу меньше числа классов, т.к. АпоВ одного из классов будет вычисляться как дополнение до единицы суммы АпоВ остальных классов.

Затем для заданной точки обучающей или тестовой выборки либо методом интер-экстраполяции, либо методом потенциальных функций (ядерного сглаживания) вычисляется АпоВ класса. Расстоянием от заданной точки до области нулевых значений той или иной АДФ серии служит модуль АДФ в заданной точке [6].

Достоинство метода в том, что он может применяться при больших объемах выборки. Размерность решаемой задачи на этапе построения серии – взвешенного МНК определяется размерностью вектор-функции $\varphi(x)$. Расчет АпоВ по серии АДФ не связан с громоздкими вычислениями.

3.2. Оценка АпоВ класса индивидуальным методом

В отличие от предыдущего в индивидуальном методе [7] расчет АпоВ класса в заданной точке основан на подборе такой АДФ, которая в заданной точке принимает нулевое или близкое к нулю значение. Параметр p^*, соответствующий подобранной АДФ и будет АпоВ класса в заданной точке. Если классов больше двух, то количество повторений процедуры будет на единицу меньше количества классов. В отличие от предыдущего, метод работает медленнее, но он не имеет предварительного этапа, его следует применять на малых объемах выборок и в случаях нестационарности. Обучающая выборка может произвольно изменяться, вычислительная часть при этом остается неизменной.

3.3. Примеры

Приведем два примера решения задач с реальными данными. Один пример интересен сравнительно большим количеством признаков. Пример взят из конкурсной задачи молодежной школы ТМШ-2014. Выборка – 252 строки. Признаков 216 – это частоты триграмм медицинских анализов. Классов два: больные и здоровые.

Произвольно по коэффициентам корреляции со столбцом с номерами классов и коэффициентам корреляции между признаками отобраны три признака. По ним получена АДФ в виде полинома второго порядка. Ошибка классификации составила 4,8%. Для линейной АДФ – 6,8%.

Во втором примере в задаче дифференциальной диагностики заболеваний с тремя классами в выборке было всего 36 строк. Были получены индивидуальным методом оценки АпоВ классов. Из 6 признаков для АДФ использованы 2. АДФ – полином второго порядка. Классификация выполнялась по классу с максимальной АопВ. В случае, если точка для расчета АпоВ классов не исключалась при расчете из выборки, получены 2 ошибки. Если исключалась, то получено 5 ошибок классификации.

4. Заключение

Свойства дискриминатной функция Андерсона позволяют по обучающей выборке с учителем непосредственно оценивать апостериорные вероятности классов. По ним не только решается задача классификации в среднем, но и объясняются принимаемые решения по каждому объекту классификации, в том числе, при различных стоимостях ошибок классификации. Отпадает нужда в надстройках типа калибратора Платта.

Список литературы

1. Platt J.C. Probabilistic Outputs for Support Vector Machines and Comparisons to Regularized Likelihood Methods. https://www.researchgate.net/publication/2594015_Probabilistic_Outputs_for_Support_Vector_Machines_and_Comparisons_to_Regularized_Likelihood_Methods.
2. Anderson T.W. An Introduction to Multivariate Statistical Analysis / 3rd ed. N.Y.: John Wiley & Sons, Inc., 2003. 711. p.
3. Зенков В.В. Использование взвешенного метода наименьших квадратов при аппроксимации дискриминантной функции цилиндрической поверхностью в задачах классификации // Автоматика и телемеханика. 2017. № 9. С. 145-158.
4. Zenkov V.V. Using Weighted Least Squares to Approximate the Discriminant Function with a Cylindrical Surface in Classification Problems // Automation and Remote Control. 2017. Vol. 78, No.9. P. 1662-1673.
5. Зенков В.В. Оценка апостериорной вероятности класса по серии дискриминантных функций Андерсона // Автоматика и телемеханика. 2019. № 3. С. 68-71.
6. Зенков В.В. Оценка вероятности принадлежности точки классу по аппроксимации одной дискриминантной функции // Автоматика и телемеханика. 2018. № 9. С. 46-58.
7. Zenkov V.V. Estimating the Probability of a Class at a Point by the Approximation of one Discriminant Function // Automation and Remote Control. 2018. Vol. 79, No. 9. P. 1580-1590.