

Лекция 4

Статистические методы распознавания,
Распознавание при заданной точности для некоторых классов,
ROC-анализ

Лектор – Сенько Олег Валентинович

Курс «Математические основы теории прогнозирования»
4-й курс, III поток

- 1 Методы, основанные на формуле Байеса
- 2 Линейный дискриминант Фишера
- 3 Логистическая регрессия
- 4 К ближайших соседей
- 5 Распознавание при заданной точности для некоторых классов

Ранее было показано, что максимальную точность распознавания обеспечивает байесовское решающее правило, относящее распознаваемый объект, описываемый вектором \mathbf{x} переменных (признаков) X_1, \dots, X_n к классу K_* , для которого условная вероятность $P(K_* | \mathbf{x})$ максимальна. Байесовские методы обучения основаны на аппроксимации условных вероятностей классов в точках признакового пространства с использованием формулы Байеса. Рассмотрим задачу распознавания классов K_1, \dots, K_L . Формула Байеса позволяет рассчитать Условные вероятности классов в точке признакового пространства могут быть рассчитаны с использованием формулы Байеса. В случае, если переменные X_1, \dots, X_n являются дискретными формула Байеса может быть записана в виде:

$$P(K_i | \mathbf{x}) = \frac{P(\mathbf{x} | K_i)P(K_i)}{\sum_{i=1}^L P(K_i)P(\mathbf{x} | K_i)} \quad (1)$$

где $P(K_1), \dots, P(K_L)$ - вероятность классов K_1, \dots, K_L безотносительно к признаковым описаниям (априорная вероятность).

В качестве оценок априорных вероятностей

$$P(K_1), \dots, P(K_L)$$

могут быть взяты доли объектов соответствующих классов в обучающей выборке. Условные вероятности $P(\mathbf{x} | K_1), \dots, P(\mathbf{x} | K_L)$ могут оцениваться на основании сделанных предположений. Например, может быть использовано предположение о независимости переменных для каждого из классов. В последнем случае вероятность $P(\mathbf{x}_j | K_i)$ для вектора $\mathbf{x}_k = (x_{j1}, \dots, x_{jn})$ может быть представлена в виде:

$$P(\mathbf{x}_j | K_i) = \prod_{i=1}^n P(X_j = x_{ki} | K_i). \quad (2)$$

Предположим, переменная X_j принимает значения из конечного множества \widetilde{M}_j^i на объектах из класса K_i при $j = 1, \dots, n$ и $i = 1, \dots, L$. Предположим, что

$$\widetilde{M}_j^i = \{a_{ji}^1, \dots, a_{ji}^{r(i,j)}\}$$

Для того, чтобы воспользоваться формулой (2) достаточно знать вероятность выполнения равенства $X_j = a_{ji}^k$ для произвольного класса и произвольной переменной. Для оценки вероятности $P(X_j = a_{ji}^k | K_i)$ может использоваться доля объектов из $\tilde{S}_t \cap K_i$, для которых $X_j = a_{ji}^k$. В случае, если переменные X_1, \dots, X_n являются непрерывными, формула Байеса может быть записана с использованием

$$P(K_i | \mathbf{x}) = \frac{p_i(\mathbf{x})P(K_i)}{\sum_{i=1}^L P(K_i)p_i(\mathbf{x})}, \quad (3)$$

где $p_1(\mathbf{x}), \dots, p_L(\mathbf{x})$ - значения плотностей вероятностей классов K_1, \dots, K_L в пространстве \mathbb{R}^n .

лотности вероятностей

$$p_1(\mathbf{x}), \dots, p_L(\mathbf{x})$$

также могут оцениваться исходя из предположения взаимной независимости переменных X_1, \dots, X_n .

В этом случае $p_i(\mathbf{x})$ может быть представлена в виде произведения одномерных плотностей

$$p_i(\mathbf{x}) = \prod_{j=1}^n p_{ji}(X_j),$$

где $p_{ji}(X_j)$ - плотность распределения переменной X_j для класса K_i . Плотности $p_{ji}(X_j)$ могут оцениваться в рамках предположения о типе распределения. Например, может использоваться гипотеза о нормальности распределений

$$p_{ji}(X_j) = \frac{1}{\sqrt{2\pi D_{ji}}} e^{-\frac{(X_j - M_{ji})^2}{2D_{ji}}},$$

где M_{ji}, D_{ji} являются математическим ожиданием и дисперсией переменной X_j . Данные параметры легко оцениваются по \tilde{S}_t .

Методы распознавания, основанные на использовании формулы Байеса в форме (1) и (3) и гипотезе о независимости переменных обычно называют **наивными байесовскими классификаторами**. Отметим, что знаменатели в правых частях формул (1) и (3) тождественны для всех классов. Поэтому при решении задач распознавания достаточно использовать только числители.

Аппроксимация плотности с помощью многомерного нормального распределения

При решении задач распознавания с помощью формулы Байеса в форме (3) могут использоваться плотности вероятности $p_1(\mathbf{x}), \dots, p_L(\mathbf{x})$, в которых переменные X_1, \dots, X_n не обязательно являются независимыми. Чаще всего используется многомерное нормальное распределение. Плотность данного распределения в общем виде представляется выражением

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma|^{\frac{1}{2}}} \exp\left[-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})\Sigma^{-1}(\mathbf{x} - \boldsymbol{\mu})^t\right], \quad (4)$$

где

$\boldsymbol{\mu}$ - математическое ожидание вектора признаков \mathbf{x} ; Σ - матрица ковариаций признаков X_1, \dots, X_n ; $|\Sigma|$ - детерминант матрицы Σ .

Для построения распознающего алгоритма достаточно оценить вектора математических ожиданий $\boldsymbol{\mu}_1, \dots, \boldsymbol{\mu}_L$ и матрицы ковариаций $\Sigma_1, \dots, \Sigma_L$ для классов K_1, \dots, K_L , соответственно.

Аппроксимация плотности с помощью многомерного нормального распределения

Оценка вектора математических ожиданий μ_i вычисляется как среднее значение векторов признаков по объектам обучающей выборки \tilde{S}_t из класса K_i :

$$\hat{\mu}_i = \frac{1}{m_i} \sum_{s_j \in \tilde{S}_t \cap K_i} \mathbf{x}_j$$

, где m_i - число объектов класса K_i в обучающей выборке. Элемент матрицы ковариаций для класса K_i вычисляется по формуле

$$\hat{\sigma}_{kk'}^i = \frac{1}{m_i} \sum_{s_j \in \tilde{S}_t \cap K_i} (x_{jk} - \mu_{ik})(x_{jk'} - \mu_{ik'}),$$

где $x_{jk} - \mu_{ik}$ - k -я компонента вектора μ_i . Матрицу ковариации, состоящую из элементов $\hat{\sigma}_{kk'}^i$ обозначим $\hat{\Sigma}_i$. Очевидно, что согласно формуле Байеса максимум $P(K_i | \mathbf{x})$ достигается для тех же самых классов для которых максимально произведение $P(K_i)p_i(\mathbf{x})$.

Использование формулы Байеса. Многомерное нормальное распределение

Очевидно, что для байесовской классификации может использоваться также натуральный логарифм $\ln[P(K_i)p_i(\mathbf{x})]$ который согласно вышеизложенному может быть оценён выражением

$$g_i(\mathbf{x}) = -\frac{1}{2}\mathbf{x}\hat{\Sigma}_i^{-1}\mathbf{x}^t + \mathbf{w}_i\mathbf{x}^t + g_i^0,$$

где $\mathbf{w}_i = \hat{\boldsymbol{\mu}}_i\hat{\Sigma}_i^{-1} g_i^0$ - не зависящее от \mathbf{x} слагаемое:
 ν_i - доля объектов класса K_i в обучающей выборке. Слагаемое g_i^0 имеет вид

$$g_i^0 = -\frac{1}{2}\hat{\boldsymbol{\mu}}_i\hat{\Sigma}_i^{-1}\hat{\boldsymbol{\mu}}_i^t - \frac{1}{2}\ln(|\hat{\Sigma}_i|) + \ln(\nu_i) - \frac{n}{2}\ln(2\pi).$$

Использование формулы Байеса. Многомерное нормальное распределение

Таким образом объект с признаковым описанием x будет отнесён построенной выше аппроксимацией байесовского классификатора к классу, для которого оценка $g_i(x)$ является максимальной. Следует отметить, что построенный классификатор в общем случае является квадратичным по признакам. Однако классификатор превращается в линейный, если оценки ковариационных матриц разных классов оказываются равными.

Использование формулы Байеса. Многомерное нормальное распределение

Рассмотрим вариант метода Линейный дискриминант Фишера (ЛДФ) для распознавания двух классов K_1 и K_2 . В основе метода лежит поиск в многомерном признаковом пространстве такого направления w , чтобы средние значения проекции на него объектов обучающей выборки из классов K_1 и K_2 максимально различались. Проекцией произвольного вектора x на направление w является отношение

$$\frac{(wx^t)}{|w|}.$$

В качестве меры различий проекций классов на используется функционал

$$\Phi(w, \tilde{S}_t) = \frac{(\hat{X}_{w1} - \hat{X}_{w2})^2}{\hat{d}_{w1} + \hat{d}_{w2}},$$

где

$$\hat{X}_{wi} = \frac{1}{m_i} \sum_{s_j \in \tilde{S}_t \cap K_i} \frac{(w x_j^t)}{|w|}$$

- среднее значение проекции векторов переменных X_1, \dots, X_n , описывающих объекты из класса K_i ;

$$\hat{d}_{wi} = \frac{1}{m_i} \sum_{s_j \in \tilde{S}_t \cap K_i} \left[\frac{(w x_j^t)}{|w|} - \hat{X}_{wi} \right]^2$$

- дисперсия проекций векторов, описывающих объекты из класса $K_i, i \in \{1, 2\}$. Смысл функционала $\Phi(w, \tilde{S}_t)$ ясен из его структуры. Он является по сути квадратом отличия между средними значениями проекций классов на направление w , нормированным на сумму внутриклассовых выборочных дисперсий.

Можно показать, что $\Phi(\mathbf{w}, \tilde{S}_t)$ достигает максимума при

$$\mathbf{w} = \hat{\Sigma}_{12}^{-1}(\hat{\boldsymbol{\mu}}_1^t - \hat{\boldsymbol{\mu}}_2^t), \quad (5)$$

где $\hat{\Sigma}_{12} = \hat{\Sigma}_1 + \hat{\Sigma}_2$. Таким образом оценка направления, оптимального для распознавания K_1 и K_2 может быть записана в виде (5)

Распознавание нового объекта s_* по векторному описанию \mathbf{x}_* производится по величине его проекции на направление \mathbf{w} :

$$\gamma(\mathbf{x}_*) = \frac{(\mathbf{w}, \mathbf{x}_*^t)}{|\mathbf{w}|}. \quad (6)$$

При этом используется простое пороговое правило: при $\gamma(\mathbf{x}_*) > b$ объект s_* относится к классу K_1 и s_* относится к классу K_2 в противном случае.

Граничный параметр b подбирается по обучающей выборке таким образом, чтобы проекции объектов разных классов на оптимальное направление w оказались бы максимально разделёнными. Простой, но эффективной, стратегией является выбор в качестве порогового параметра b средней проекции объектов обучающей выборки на w . Метод ЛДФ легко обобщается на случай с несколькими классами. При этом исходная задача распознавания классов K_1, \dots, K_L сводится к последовательности задач с двумя классами K'_1 и K'_2 :

Зад. 1. Класс $K'_1 = K_1$, класс $K'_2 = \Omega \setminus K_1$

.....

Зад. L. Класс $K'_1 = K_L$, класс $K'_2 = \Omega \setminus K_L$

Для каждой из L задач ищется оптимальное направление. В результате получается набор из L направлений w_1, \dots, w_L .

В результате получается набор из L направлений w_1, \dots, w_L . При распознавании нового объекта s_* по признаковому описанию x_* вычисляются проекции на w_1, \dots, w_L :

$$\gamma_1(x_*) = \frac{(w_1, x_*^t)}{|w_1|}, \dots, \gamma_L(x_*) = \frac{(w_L, x_*^t)}{|w_L|}.$$

Распознаваемый объект относится к тому классу, соответствующему максимальной величине проекции. Распознавание может производиться также по величинам

$$[\gamma_1(x_*) - b_1], \dots, [\gamma_L(x_*) - b_L].$$

Логистическая регрессия.

Целью логистической регрессии является аппроксимация плотности условных вероятностей классов в точках признакового пространства. При этом аппроксимация производится с использованием логистической функции.

$$g(z) = \frac{e^z}{1 + e^z} = \frac{1}{1 + e^{-z}}$$

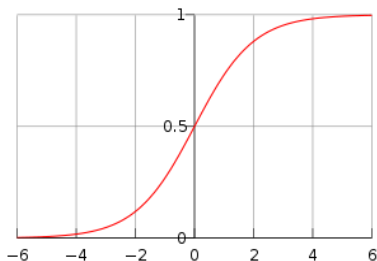


Рис 1. Логистическая функция.

В методе логистическая регрессия связь условной вероятности класса K с прогностическими признаками осуществляются через переменную Z , которая задаётся как линейная комбинация признаков:

$$z = \beta_0 + \beta_1 X_1 + \dots + \beta_n X_n.$$

Условная вероятность K в точке векторного пространства $\mathbf{x}_* = (x_{1*}, \dots, x_{n*})$ задаётся в виде

$$P(K | \mathbf{x}) = \frac{e^{\beta_0 + \beta_1 X_1 + \dots + \beta_n X_n}}{1 + e^{\beta_0 + \beta_1 X_1 + \dots + \beta_n X_n}} = \frac{1}{1 + e^{-\beta_0 - \beta_1 X_1 - \dots - \beta_n X_n}} \quad (7)$$

Оценки регрессионных параметров $\beta_0, \beta_1, \dots, \beta_n$ могут быть вычислены по обучающей выборке с помощью различных вариантов метода максимального правдоподобия. Предположим, что объекты обучающей выборки сосредоточены в точках признакового пространства из множества $\tilde{\mathbf{x}} = \{\mathbf{x}_1, \dots, \mathbf{x}_r\}$. При этом распределение объектов обучающей выборки по точкам задаётся с помощью набора пар $\{(m_1, k_1), \dots, (m_r, k_r)\}$, где m_i - общее число объектов в точке \mathbf{x}_i , k_i - число объектов класса K в точке \mathbf{x}_i . Вероятность данной конфигурации подчиняется распределению Бернулли. Введём обозначение $\varrho(\mathbf{x}) = P(K | \mathbf{x})$. Оценка вектора регрессионных параметров $\beta = (\beta_0, \dots, \beta_n)$ может быть получена с помощью метода максимального правдоподобия. Функция правдоподобия может быть записана в виде

$$L(\beta, \tilde{\mathbf{x}}) = \prod_{j=1}^r C_{m_i}^{k_i} [\varrho(\mathbf{x})_j]^{k_j} [1 - \varrho(\mathbf{x})_j]^{(m_j - k_j)} \quad (8)$$

Принимая во внимание справедливость равенств

$$\frac{\varrho(\mathbf{x})}{1 - \varrho(\mathbf{x})} = e^{\beta_0 + \beta_1 X_1 + \dots + \beta_n X_n},$$

$$1 - \varrho(\mathbf{x}) = \frac{1}{1 + e^{\beta_0 + \beta_1 X_1 + \dots + \beta_n X_n}},$$

приходим равенству

$$L(\boldsymbol{\beta}, \tilde{\mathbf{x}}) = \prod_{j=1}^r C_{m_i}^{k_i} e^{k_i \beta_0 + \beta_1 x_{j1} + \dots + \beta_n x_{jn}} \frac{1}{(1 + e^{\beta_0 + \beta_1 x_{j1} + \dots + \beta_n x_{jn}})^{m_i}} \quad (9)$$

Поиск оптимального значения параметров удобнее производить, решая задачу максимизации логарифма функции правдоподобия, который в нашем случае принимает вид:

$$\begin{aligned}\ln[L(\boldsymbol{\beta}, \tilde{\boldsymbol{x}})] &= \sum_{j=1}^r \ln C_{m_j}^{k_j} + \sum_{j=1}^r [k_j(\beta_0 + \beta_1 x_{j1} + \dots + \beta_n x_{jn})] + \\ &+ \sum_{j=1}^r m_j \ln\left(\frac{1}{1 + e^{\beta_0 + \beta_1 x_{j1} + \dots + \beta_n x_{jn}}}\right)\end{aligned}$$

Простым, но достаточно эффективным подходом к решению задач распознавания является метод k -ближайших соседей. Оценка условных вероятностей $P(K_i | \mathbf{x})$ ведётся по ближайшей окрестности V_k точки \mathbf{x} , содержащей k признаковых описаний объектов обучающей выборки. В качестве оценки выступает отношение $\frac{k_i}{k}$, где k_i - число признаковых описаний объектов обучающей выборки из K_i внутри V_k . Окрестность V_k задаётся с помощью функции расстояния $\rho(\mathbf{x}', \mathbf{x}'')$ заданной на декартовом произведении $\tilde{X} \times \tilde{X}$, где \tilde{X} - область допустимых значений признаковых описаний. В качестве функции расстояния может быть использована стандартная евклидова метрика. То есть расстояние между двумя векторами $\mathbf{x}' = (x'_1, \dots, x'_n)$ и $\mathbf{x}'' = (x''_1, \dots, x''_n)$

$$\rho(\mathbf{x}', \mathbf{x}'') = \sqrt{\frac{1}{n} \sum_{i=1}^n (x'_i - x''_i)^2}.$$

Для задач с бинарными признаками в качестве функции расстояния может быть использована метрика Хэмминга, равная числу совпадающих позиций в двух сравниваемых признаковых описаниях. Окрестность V_k ищется путём поиска в обучающей выборке \tilde{S}_t векторных описаний, ближайших в смысле выбранной функции расстояний, к описанию x_* распознаваемого объекта s_* . Единственным параметром, который может быть использован для настройки (обучения) алгоритмов в методе k -ближайших соседей является собственно само число ближайших соседей. Для оптимизации параметра k обычно используется метод, основанный на скользящем контроле. Оценка точности распознавания производится по обучающей выборке при различных k и выбирается значение данного параметра, при котором полученная точность максимальна.

Байесовский классификатор обеспечивает максимальную общую точность распознавания. Однако при решении конкретных практических задач потери, связанные с неправильной классификацией объектов, принадлежащих к одному из классов, значительно превышают потери, связанные с неправильной классификацией объектов других классов. Для оптимизации потерь необходимо использование методов распознавания с учётом предпочтительной точности распознавания для некоторых классов. Одним из возможных подходов является фиксирование порога для точности распознавания одного из классов. Оптимальное решающее правило в задаче распознавания с двумя классами K_1 и K_2 , обеспечивающее максимальную точность распознавания K_2 при фиксированной точности распознавания K_1 , описывается критерием Неймана-Пирсона.

Критерий Неймана-Пирсона Максимальная точность распознавания K_2 при точности распознавания K_1 равной α обеспечивается правилом: Объект с описанием \mathbf{x} относится в класс K_1 , если

$$P(K_1 | \mathbf{x}) \geq \eta P(K_2 | \mathbf{x})$$

где параметр η определяется из условия

$$\int_{\Theta} P(K_1 | \mathbf{x}) p(\mathbf{x}) d\mathbf{x} = \alpha$$

$$\Theta = \{\mathbf{x} | P(K_1 | \mathbf{x}) \geq \eta P(K_2 | \mathbf{x})\}$$

$p(\mathbf{x})$ - плотность распределения $K_1 \cup K_2$ в точке \mathbf{x} . Критерий Неймана-Пирсона может быть использован, если известны плотности распределения распознаваемых классов. Плотности могут быть восстановлены в рамках Байесовских методов обучения на основе гипотез о виде распределений. ,

Распознавание при заданной точности распознавания некоторых классов

Критерий Неймана-Пирсона может быть использован, если известны плотности распределения распознаваемых классов. Плотности могут быть восстановлены в рамках Байесовских методов обучения на основе гипотез о виде распределений. Однако существуют эффективные средства регулирования точности распознавания при предпочтительности одного из классов, которые не требуют гипотез о виде распределения. Данные средства основаны на структуре распознающего алгоритма. Каждый алгоритм распознавания классов K_1, \dots, K_l может быть представлен как последовательное выполнение распознающего оператора \mathbf{R} и решающего правила :

$$A = \mathbf{R} \otimes C.$$

Оператор оценок вычисляет для распознаваемого объекта s вещественные оценки $\gamma_1, \dots, \gamma_L$ за классы K_1, \dots, K_l соответственно.

Распознавание при заданной точности распознавания некоторых классов

Решающее правило производит отнесение объекта s по вектору оценок $\gamma_1, \dots, \gamma_L$ к одному из классов. Распространённым решающим правилом является простая процедура, относящая объект в тот класс, оценка за который максимальна.

В случае распознавания двух классов K_1 и K_2 распознаваемый объект s будет отнесён к классу K_1 , если $\gamma_1(s) - \gamma_2(s) > 0$ и классу K_2 в противном случае. Назовём приведённое выше правило правилом $C(0)$. Однако точность распознавания правила $C(0)$ может оказаться слишком низкой для того, чтобы обеспечить требуемую величину потерь, связанных с неправильной классификацией объектов, на самом деле принадлежащих классу K_1 . Для достижения необходимой величины потерь может быть использовано пороговое решающее правило $C(\delta)$.

Распознавание при заданной точности распознавания некоторых классов

Правило $\mathbf{C}(\delta)$: распознаваемый объект s будет отнесён к классу K_1 , если $\gamma_1(s) - \gamma_2(s) > \delta$ и классу K_2 в противном случае. Обозначим через $p_{ci}(\delta, s)$ вероятность правильной классификации правилом $\mathbf{C}(\delta)$ объекта s , на самом деле принадлежащего K_i , $i \in \{1, 2\}$. При $\delta < 0$ $p_{c1}(\delta, s) \geq p_{c1}(0, s)$, но $p_{c2}(\delta, s) \leq p_{c2}(0, s)$. Уменьшая δ , мы увеличиваем $p_{c1}(\delta, s)$ и уменьшаем $p_{c2}(\delta, s)$. Напротив, увеличивая δ , мы уменьшаем $p_{c1}(\delta, s)$ и увеличиваем $p_{c2}(\delta, s)$. Зависимость между $p_{c1}(\delta, s)$ и $p_{c2}(\delta, s)$ может быть приближённо восстановлена по обучающей выборке \tilde{S}_t , включающей описания объектов $\{s_1, \dots, s_m\}$.

Пусть

$$\begin{pmatrix} \gamma_1(s_1) & \dots & \gamma_1(s_m) \\ \gamma_2(s_1) & \dots & \gamma_2(s_m) \end{pmatrix}$$

является матрицей оценок за классы K_1 и K_2 объектов из \tilde{S}_t . Пусть

$$\gamma(s_1) = \gamma_1(s_1) - \gamma_2(s_1), \dots, \gamma(s_m) = \gamma_1(s_m) - \gamma_2(s_m).$$

Предположим, что величины $[\gamma(s_1), \dots, \gamma(s_m)]$ принимают r значений $\Gamma_1, \dots, \Gamma_r$. Рассмотрим r пороговых решающих правил $[C(\Gamma_1), \dots, C(\Gamma_r)]$. Для каждого из правил $C(\Gamma_i)$ обозначим через $\nu_{c1}(\Gamma_i)$ долю K_1 среди объектов обучающей выборки, удовлетворяющих условию $\gamma(s_*) \geq \Gamma_i$, а через $\nu_{c2}(\Gamma_i)$ обозначим долю K_2 среди объектов обучающей выборки, удовлетворяющих условию $\gamma(s_*) < \Gamma_i$.

Отообразим результаты расчётов

$$\{[\nu_{c1}(\Gamma_1), \nu_{c2}(\Gamma_1)] \dots, [\nu_{c1}(\Gamma_r), \nu_{c2}(\Gamma_r)]\}$$

как точки на в декартовой системе координат. Соединив точки отрезками прямых, получим ломаную линию (I), соединяющую точки (1,0) и (0,1). Данная линия графически отображает аппроксимацию по обучающей выборке взаимозависимости между $p_{c1}(\delta, s)$ и $p_{c2}(\delta, s)$ при всевозможных значениях δ . Соответствующий пример представлен на рисунке 2.

Распознавание при заданной точности распознавания некоторых классов. ROC анализ.

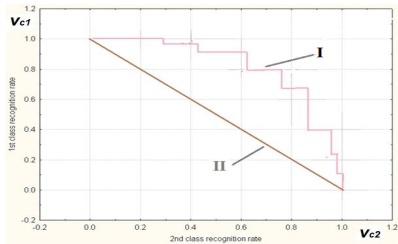


Рис.2

Взаимозависимость между ν_{c1} и ν_{c2} наиболее полно оценивает эффективность распознающего оператора \mathbf{R} . Отметим, что ν_{c1} постепенно убывает по мере роста ν_{c2} . .

Распознавание при заданной точности распознавания некоторых классов. ROC анализ.

Однако сохранение высокого значения ν_{c1} при высоких значениях ν_{c2} соответствует существованию решающего правила, при котором точность распознавания обоих классов высока. Таким образом эффективному распознающему оператору соответствует близость линии I к прямой, связывающей точки (0,1) и (1,1). То есть наиболее высокой эффективности соответствует максимально большая площадь под линией I. Отсутствию распознающей способности соответствует близость к прямой II, связывающей точки (0, 1) и (1,0). На рисунке 3 сравниваются линии, характеризующие эффективность распознающих операторов, принадлежащих к трём методам распознавания, при решении задачи распознавания двух видов аутизма по психометрическим показателям .

Распознавание при заданной точности распознавания некоторых классов. ROC анализ.

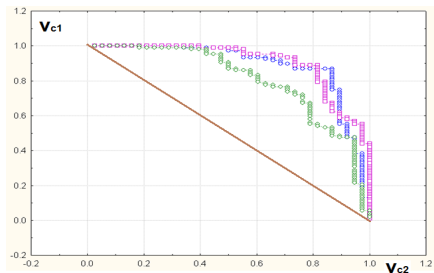





Рис.3

-  - линейный дискриминант Фишера;
-  - метод опорных вектор;
-  - статистически взвешенные синдромы.

Методы распознавания используются при решении многих задач идентификации объектов, представляющих важность для пользователя. Эффективность идентификации для таких задач удобно описывать в терминах: «Чувствительность» - доля правильно распознанных объектов целевого класса «Ложная тревога» - доля объектов ошибочно отнесённых в целевой класс. Пример кривой, связывающей параметры «Чувствительность» и «Ложная тревога» представлен на рисунке 4. Анализ, основанный на построении и анализе линий, связывающих параметры «Чувствительность» и «Ложная тревога» принято называть анализом Receiver Operating Characteristic или ROC-анализом. Линии, связывающих параметры «Чувствительность» и «Ложная тревога» принято называть ROC-кривыми.

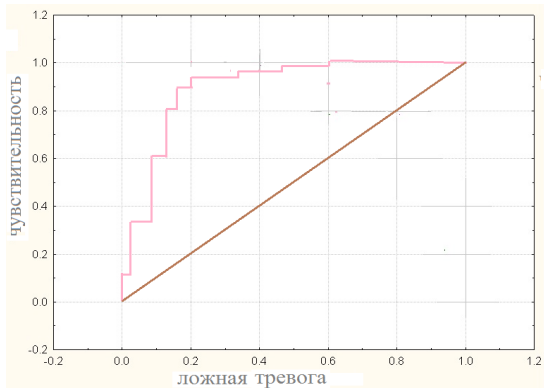


Рис.4. Пример ROC кривой