

Задачи классификации и регрессии

Технологии обработки данных

- ***Knowledge Discovery in Databases (KDD)*** – процесс получения из данных знаний в виде зависимостей, правил, моделей.
- ***Data Mining (DM)*** – обнаружение в данных неизвестных нетривиальных знаний, необходимых для принятия решений.

Методы ДМ

- **Классификация** – установление зависимости дискретной выходной переменной от входных.
- **Кластеризация** – группировка объектов на основе свойств.
- **Регрессия** - установление зависимости непрерывной выходной переменной от входных.
- **Ассоциация** – выявление закономерностей между связанными событиями.
- **Последовательные шаблоны** – установление закономерностей между связанными во времени событиями

Решение задач классификации

- *Логистическая регрессия*
- *Метод опорных векторов*
- *Деревья решений*
- *Байесовские алгоритмы*

Задача классификации данных

Задача классификации формулируется следующим образом.

Имеется множество объектов

$$X = \{\mathbf{x}^h\}, h = \overline{1, N}$$

Каждый объект характеризуется набором (x_1, \dots, x_n)

свойств и меткой принадлежности к q

классу из множества $q \in Q$

классов при известном наборе \mathbf{x}

характеристик некоторый объект необходимо отнести к какому-либо $q \in Q$ классу

Свойства объекта могут быть булевыми, дискретными или непрерывными, а метку класса обычно представляют в виде номера класса, то есть дискретной переменной.

Метод опорных векторов

Рассмотрим задачу бинарной классификации. Имеющийся набор данных

$$X = \{x^h, q^h\}, h = \overline{1, H}$$

содержит два
класса

$$q \in \{+1; -1\}$$

Требуется построить поверхность, разделяющую все множество точек на два подмножества

$$q = 1 \quad \text{и} \quad q = -1$$

Сначала предположим возможность линейного разделения, то есть, поверхность представляет собой разделительную гиперплоскость.

Уравнение разделительной гиперплоскости в пространстве переменных

представим в виде $x \in R^n$

$$\sum_{j=1}^n w_j x_j + b = \mathbf{w}^T \mathbf{x} + b = 0$$

а линейный пороговый классификатор $q^{SL}(\mathbf{x}) = \text{sign}(\mathbf{w}^T \mathbf{x} + b)$

Коэффициенты $w_j, j = \overline{1, n}; b$

подбираются в процессе обучения.

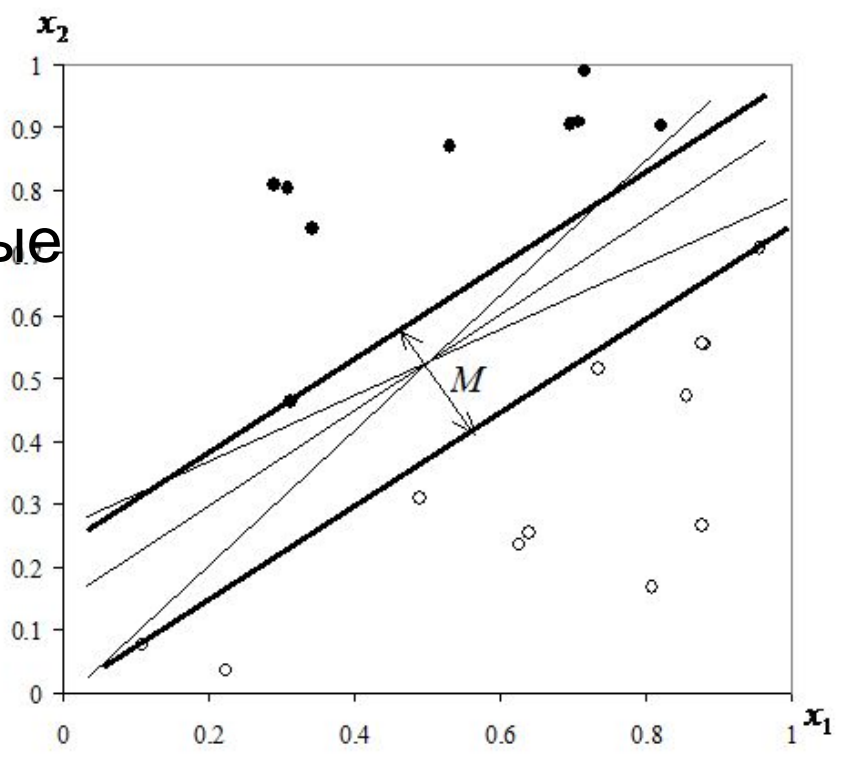
На рис. представлены два класса в двумерном пространстве.

Класс

$q = 1$

черные

точки



$q = -1$
светлые
точки.

Видно, что прямая линия, разделяющая две группы точек, не является единственной. Наиболее уверенная классификация обеспечивается, если точки разных классов максимально далеко находятся от разделительной линии. Это будет выполнено, если ширина M пустой полосы между точками разных классов будет максимальна.

Коэффициент
ы w_j, b

можно нормировать таким
образом, чтобы $w^T x + b = q$

в точках, ближайших к разделяющей классы полосе. В
остальных точках условие $w^T x + b \begin{cases} \geq 1 & | q = 1 \\ \leq -1 & | q = -1 \end{cases}$

выполняется с запасом. Ширина полосы M рассчитывается
как расстояние между двумя параллельными плоскостями
 $w^T x + b = -1$ и $w^T x + b = 1$ $M = \frac{(b+1) - (b-1)}{\sqrt{w^T w}} = \frac{2}{\sqrt{w^T w}}$

Следовательно, максимум M соответствует

и получим задачу оптимизации с
ограничениями

$$q^h (w^T x^h + b) \geq 1, h = \overline{1, H}$$

Функция Лагранжа

$$L(w, b, \Lambda) = 0.5w^T w - \sum_{h=1}^H \lambda_h (q^h (w^T x^h + b) - 1),$$

где $\Lambda = (\lambda_h), h = \overline{1, H}$ - множители Лагранжа, $\lambda_h \geq 0$,

$$\frac{\partial L}{\partial w_j} = w_j - \sum_{h=1}^H \lambda_h q^h x_j^h = 0, \text{ или } w_j = \sum_{h=1}^H \lambda_h q^h x_j^h, j = \overline{1, n},$$

$$\frac{\partial L}{\partial b} = \sum_{h=1}^H \lambda_h q^h = 0.$$

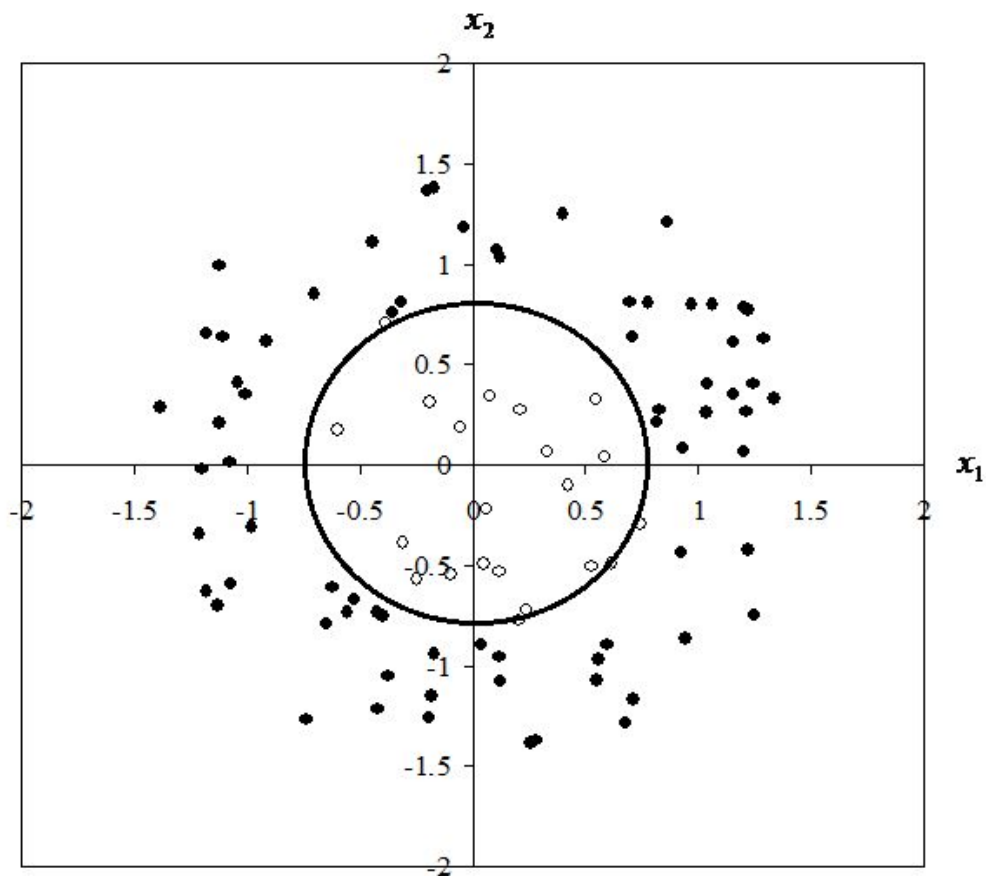
$$\Phi(\Lambda) = \sum_{h=1}^H \lambda_h - 0.5 \sum_{j=1}^H \sum_{h=1}^H \lambda_j \lambda_h q^j q^h (x^j)^T \rightarrow \max,$$

$$\sum_{h=1}^h \lambda_h q^h = 0,$$

$$\lambda_h \geq 0, h = \overline{1, H}.$$

Коэффициент b определяется как среднее по всем опорным векторам: $b = q^h - w^t x^h$.

Если в данных присутствует существенная нелинейность, то решение задачи не приводит к правильной классификации. На рисунке приведены данные такого типа.



Выражение для
классификатора

$$q^{SVM}(\mathbf{x}) = \text{sign}\left(\sum_{h=1}^H \lambda_h q^h K(\mathbf{x}^h, \mathbf{x}) + b\right)$$

Практически используются следующие функции
ядра:

$$K(\mathbf{x}, \mathbf{x}') = \mathbf{x}^T \mathbf{x}'$$

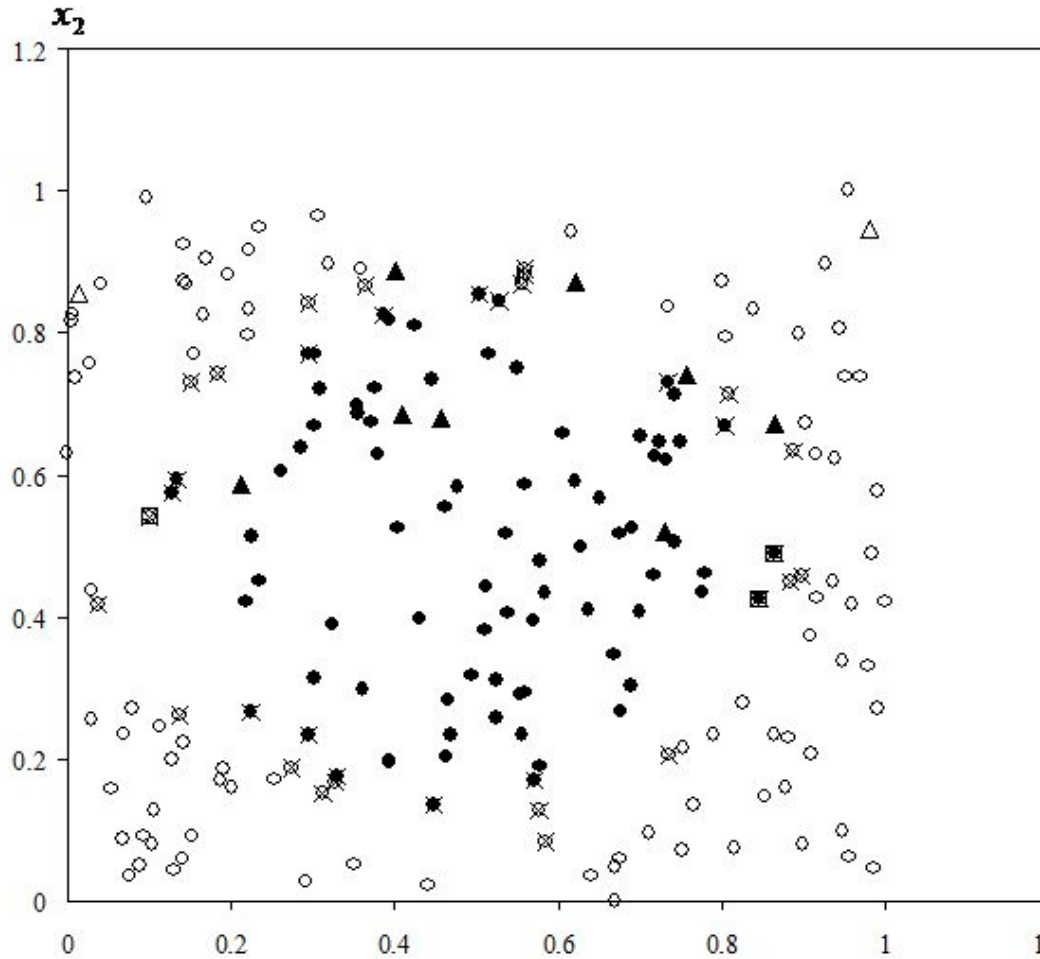
$$K(\mathbf{x}, \mathbf{x}') = (\mathbf{x}^T \mathbf{x}' + 1)^d$$

$$K(\mathbf{x}, \mathbf{x}') = \exp(-0.5(\mathbf{x} - \mathbf{x}')^T \Sigma^{-1}(\mathbf{x} - \mathbf{x}'))$$

$$K(\mathbf{x}, \mathbf{x}') = \tanh(\mathbf{x}^T \mathbf{x}' + w_0)$$

$$K(\mathbf{x}, \mathbf{x}') = \frac{1}{\sqrt{\|\mathbf{x} - \mathbf{x}'\|^2 + \beta}} \mathbf{x}^T \mathbf{x}'$$

Для примера на рис. представлена выборка из 190 точек.



Нелинейное разделение с радиальным ядром

Опорные вектора отмечены крестиками. Эти точки расположены на разделяющей поверхности замкнутой поверхности. Неправильно классифицированные точки обведены квадратами. Таких точек три из 190. Треугольниками обозначены точки из проверочной выборки, не участвующие в обучении.

Байесовские классификаторы

Вероятностная постановка задачи классификации и построение классификаторов основывается на имеющихся прецедентах (\mathbf{x}^h, q^h) , $h = \overline{1, H}$ (обучающей выборке). Выборка представляет набор случайных величин из генеральной совокупности с плотностью распределения $p(\mathbf{x}, q)$. Обозначим вероятность появления объекта каждого класса через $P_q = p(q)$. Это априорные вероятности. Плотность распределения классов называется функцией правдоподобия классов $p_q(\mathbf{x}) = p(\mathbf{x} | q)$. Плотность распределения вероятностного пространства объектов $p(\mathbf{x}, q) = p(q)p(\mathbf{x} | q)$. Функция плотности $p(\mathbf{x}, q)$ неизвестна и основной проблемой является ее оценка на имеющейся обучающей выборке.

При известных априорных вероятностях и функциях правдоподобия байесовский классификатор представляется в виде

$$q^B(\mathbf{x}) = \arg \max p(q)p(\mathbf{x} | q). \quad (1)$$

В соответствии с определением условной вероятности ()

$$p(q)p(\mathbf{x} | q) = p(q | \mathbf{x})p(\mathbf{x}) = p(q | \mathbf{x}) \sum_{r \in Q} p_r(\mathbf{x})p(r).$$

Отсюда определяется величина апостериорной вероятности

$$p(q | \mathbf{x}) = \frac{p(q)p(\mathbf{x} | q)}{\sum_{r \in Q} p_r(\mathbf{x})p(r)}. \quad (2)$$

Апостериорная вероятность показывает вероятность принадлежности объекта \mathbf{x} к каждому классу $q \in Q$. Принцип максимума апостериорной вероятности определяет номер класса, к которому относится рассматриваемый объект при равнозначности всех классов. Формулу (1) можно записать с применением формулы (2)

$$q^B(\mathbf{x}) = \arg \max p(q | \mathbf{x}). \quad (3)$$

При известной плотности распределения $p(\mathbf{x}, q)$ классификатор (1) или (3) является оптимальным. Для определения априорных вероятностей трудности нет. Эти вероятности оцениваются по частоте появления объектов каждого класса. При увеличении мощности выборки $|X| = N \rightarrow \infty$

частота $\frac{N_q}{N}$ сходится к P_q . Способы восстановления неизвестной

плотности $p_q(\mathbf{x}) = p(\mathbf{x} | q)$ определяют различные разновидности байесовских классификаторов. Ряд упрощающих предположений облегчает обучение классификаторов.

Наивный байесовский классификатор

Упрощенный (наивный) алгоритм Байеса является алгоритмом классификации на основе байесовского правила в предположении, что входные переменные (свойства объекта) условно независимы друг от друга. В этом случае

$$p(\mathbf{x} | q) = \prod_{j=1}^n p(x_j | q)$$

и правило классификации для наивного байесовского классификатора принимает вид

$$q^{NB}(\mathbf{x}) = \arg \max p(q) \prod_{j=1}^n p(x_j | q), \quad (4)$$

так как знаменатель в (2) не зависит от q .

Рассмотрим пример классификации с $n = 3, H = 30$ для дискретных входных переменных. Дискретные значения входных переменных в следующей таблице

x_1	x_2	x_3	q
1	2	1	4
1	1	1	1
1	1	1	1
2	1	1	2
2	1	2	3
1	2	2	5
2	1	1	2
1	1	1	1
2	1	1	2
1	2	1	4
2	2	1	5
2	2	1	5
1	1	2	2
1	2	1	4
2	2	2	5
2	2	1	5
2	2	2	5
2	2	1	5
2	1	1	2
2	1	2	3
2	2	2	5
2	1	1	2
2	2	2	5
2	1	1	2
2	2	2	5
2	2	1	5
1	2	1	4
2	1	1	2
2	2	1	5
1	1	1	1

Входные переменные $x_j, j = 1, 2, 3$ разбиты на 2 класса ($l = 1, 2$), выходная переменная принадлежит пяти классам. Количество точек H_q , принадлежащих классу q и количество точек H_{qjl} , соответствующих классу q , номеру переменной x_j , классу входной переменной l запишем в таблицу

Таблица . Количество точек, соответствующих классам

H_{qjl}			H_q	q	l
$j=1$	$j=2$	$j=3$			
4	4	4	4	1	1
0	0	0	4	1	2
1	8	7	8	2	1
7	0	1	8	2	2
0	2	0	2	3	1
2	0	2	2	3	2
4	0	4	4	4	1
0	4	0	4	4	2
1	0	6	12	5	1
11	12	6	12	5	2

Соответствующие вероятности $p(q) = \frac{H_q}{H}$, $p(x_j | q) = \frac{H_{qj}}{H_q}$ содержатся в

следующей таблице.

Таблица. Расчет вероятностей

$p(x_j q)$					
$j=1$	$j=2$	$j=3$	$p(q)$	q	L
1	1	1	0.1333	1	1
0	0	0	0.1333	1	2
0.125	1	0.875	0.2667	2	1
0.875	0	0.125	0.2667	2	2
0	1	0	0.0667	3	1
1	0	1	0.0667	3	2
1	0	1	0.1333	4	1
0	1	0	0.1333	4	2
0.0833	0	0.5	0.4	5	1
0.9167	1	0.5	0.4	5	2

Эта таблица является результатом обучения наивного байесовского классификатора.

Рассмотрим точку $\mathbf{x} = (2;1;1)^T$. Апостериорные вероятности $p(q | \mathbf{x})$ для нее равны:

q	$p(q \mathbf{x})$		
1	0	$0.133 \cdot 0 \cdot 1 \cdot 1$	
2	0.204	$0.2667 \cdot 0.875 \cdot 1 \cdot 0.875$	max
3	0	$0.0667 \cdot 1 \cdot 1 \cdot 0$	
4	0	$0.133 \cdot 0 \cdot 0 \cdot 1$	
5	0	$0.4 \cdot 0.9167 \cdot 0 \cdot 0.5$	

Следовательно, точка $\mathbf{x} = (2;1;1)^T$ принадлежит второму классу с вероятностью 0.204.

$p(x_j q)$					
$j=1$	$j=2$	$j=3$	$p(q)$	q	L
1	1	1	0.1333	1	1
0	0	0	0.1333	1	2
0.125	1	0.875	0.2667	2	1
0.875	0	0.125	0.2667	2	2
0	1	0	0.0667	3	1
1	0	1	0.0667	3	2
1	0	1	0.1333	4	1
0	1	0	0.1333	4	2
0.0833	0	0.5	0.4	5	1
0.9167	1	0.5	0.4	5	2

Последовательность действий

1	2	1	4
1	1	1	1
1	1	1	1
2	1	1	2
2	1	2	3
1	2	2	5
2	1	1	2
1	1	1	1
2	1	1	2
1	2	1	4
2	2	1	5
2	2	1	5
1	1	2	2
1	2	1	4
2	2	2	5
2	2	1	5
2	2	2	5
2	2	1	5
2	1	1	2
2	1	2	3
2	2	2	5
2	1	1	2
2	2	2	5
2	2	1	5
1	2	1	4
2	1	1	2
2	2	1	5
1	1	1	1
2	1	1	

$$H_{jq}^h$$

j=1 j=2 j=3 Nq

4 4 4 4 q=1,l=1

1	2	1	4
1	1	1	1
1	1	1	1
2	1	1	2
2	1	2	3
1	2	2	5
2	1	1	2
1	1	1	1
2	1	1	2
1	2	1	4
2	2	1	5
2	2	1	5
1	1	2	2
1	2	1	4
2	2	2	5
2	2	1	5
2	2	2	5
2	2	1	5
2	1	1	2
2	1	2	3
2	2	2	5
2	1	1	2
2	2	2	5
2	1	1	2
2	2	2	5
2	2	1	5
1	2	1	4
2	1	1	2
2	2	1	5
1	1	1	1
2	1	1	

j=1	j=2	j=3	Nq	
4	4	4	4	q=1,l=1
0	0	0	4	q=1,l=2

1	2	1	4
1	1	1	1
1	1	1	1
2	1	1	2
2	1	2	3
1	2	2	5
2	1	1	2
1	1	1	1
2	1	1	2
1	2	1	4
2	2	1	5
2	2	1	5
1	1	2	2
1	2	1	4
2	2	2	5
2	2	1	5
2	2	2	5
2	2	1	5
2	1	1	2
2	1	2	3
2	2	2	5
2	1	1	2
2	2	2	5
2	1	1	2
2	2	2	5
2	2	1	5
1	2	1	4
2	1	1	2
2	2	1	5
1	1	1	1
2	1	1	

j=1	j=2	j=3	Nq	
4	4	4	4	q=1,l=1
0	0	0	4	q=1,l=2
1	8	7	8	q=2,l=1

1	2	1	4
1	1	1	1
1	1	1	1
2	1	1	2
2	1	2	3
1	2	2	5
2	1	1	2
1	1	1	1
2	1	1	2
1	2	1	4
2	2	1	5
2	2	1	5
1	1	2	2
1	2	1	4
2	2	2	5
2	2	1	5
2	2	2	5
2	2	1	5
2	1	1	2
2	1	2	3
2	2	2	5
2	1	1	2
2	2	2	5
2	2	1	5
1	2	1	4
2	1	1	2
2	2	1	5
1	1	1	1
2	1	1	

j=1	j=2	j=3	Nq	
4	4	4	4	q=1,l=1
0	0	0	4	q=1,l=2
1	8	7	8	q=2,l=1
7	0	1	8	q=2,l=2

1	2	1	4
1	1	1	1
1	1	1	1
2	1	1	2
2	1	2	3
1	2	2	5
2	1	1	2
1	1	1	1
2	1	1	2
1	2	1	4
2	2	1	5
2	2	1	5
1	1	2	2
1	2	1	4
2	2	2	5
2	2	1	5
2	2	2	5
2	2	1	5
2	1	1	2
2	1	2	3
2	2	2	5
2	1	1	2
2	2	2	5
2	2	1	5
1	2	1	4
2	1	1	2
2	2	1	5
1	1	1	1
2	1	1	

j=1	j=2	j=3	Nq	
4	4	4	4	q=1,l=1
0	0	0	4	q=1,l=2
1	8	7	8	q=2,l=1
7	0	1	8	q=2,l=2
0	2	0	2	q=3,l=1

1	2	1	4
1	1	1	1
1	1	1	1
2	1	1	2
2	1	2	3
1	2	2	5
2	1	1	2
1	1	1	1
2	1	1	2
1	2	1	4
2	2	1	5
2	2	1	5
1	1	2	2
1	2	1	4
2	2	2	5
2	2	1	5
2	2	2	5
2	2	1	5
2	1	1	2
2	1	2	3
2	2	2	5
2	1	1	2
2	2	2	5
2	2	1	5
1	2	1	4
2	1	1	2
2	2	1	5
1	1	1	1
2	1	1	

j=1	j=2	j=3	Nq	
4	4	4	4	q=1,l=1
0	0	0	4	q=1,l=2
1	8	7	8	q=2,l=1
7	0	1	8	q=2,l=2
0	2	0	2	q=3,l=1
2	0	2	2	q=3,l=2

1	2	1	4
1	1	1	1
1	1	1	1
2	1	1	2
2	1	2	3
1	2	2	5
2	1	1	2
1	1	1	1
2	1	1	2
1	2	1	4
2	2	1	5
2	2	1	5
1	1	2	2
1	2	1	4
2	2	2	5
2	2	1	5
2	2	2	5
2	2	1	5
2	1	1	2
2	1	2	3
2	2	2	5
2	1	1	2
2	2	2	5
2	2	1	5
1	2	1	4
2	1	1	2
2	2	1	5
1	1	1	1
2	1	1	

j=1	j=2	j=3	Nq	
4	4	4	4	q=1,l=1
0	0	0	4	q=1,l=2
1	8	7	8	q=2,l=1
7	0	1	8	q=2,l=2
0	2	0	2	q=3,l=1
2	0	2	2	q=3,l=2
4	0	4	4	q=4,l=1

1	2	1	4
1	1	1	1
1	1	1	1
2	1	1	2
2	1	2	3
1	2	2	5
2	1	1	2
1	1	1	1
2	1	1	2
1	2	1	4
2	2	1	5
2	2	1	5
1	1	2	2
1	2	1	4
2	2	2	5
2	2	1	5
2	2	2	5
2	2	1	5
2	1	1	2
2	1	2	3
2	2	2	5
2	1	1	2
2	2	2	5
2	2	1	5
1	2	1	4
2	1	1	2
2	2	1	5
1	1	1	1
2	1	1	

j=1	j=2	j=3	Nq	
4	4	4	4	q=1,l=1
0	0	0	4	q=1,l=2
1	8	7	8	q=2,l=1
7	0	1	8	q=2,l=2
0	2	0	2	q=3,l=1
2	0	2	2	q=3,l=2
4	0	4	4	q=4,l=1
0	4	0	4	q=4,l=2

1	2	1	4					
1	1	1	1					
1	1	1	1					
2	1	1	2	j=1	j=2	j=3	Nq	
2	1	2	3	4	4	4	4	q=1,l=1
1	2	2	5	0	0	0	4	q=1,l=2
2	1	1	2	1	8	7	8	q=2,l=1
1	1	1	1	7	0	1	8	q=2,l=2
2	1	1	2	0	2	0	2	q=3,l=1
1	2	1	4	2	0	2	2	q=3,l=2
2	2	1	5	4	0	4	4	q=4,l=1
2	2	1	5	0	4	0	4	q=4,l=2
1	1	2	2					
1	2	1	4	1	0	6	12	q=5,l=1
2	2	2	5					
2	2	1	5					
2	2	2	5					
2	2	1	5					
2	1	1	2					
2	1	2	3					
2	2	2	5					
2	1	1	2					
2	2	2	5					
2	2	1	5					
2	1	1	2					
2	2	2	5					
2	2	1	5					
1	2	1	4					
2	1	1	2					
2	2	1	5					
1	1	1	1					
2	1	1						

1	2	1	4						
1	1	1	1						
1	1	1	1						
2	1	1	2	j=1	j=2	j=3	Nq		
2	1	2	3	4	4	4	4	q=1,l=1	
1	2	2	5	0	0	0	4	q=1,l=2	
2	1	1	2	1	8	7	8	q=2,l=1	
1	1	1	1	7	0	1	8	q=2,l=2	
2	1	1	2	0	2	0	2	q=3,l=1	
1	2	1	4	2	0	2	2	q=3,l=2	
2	2	1	5	4	0	4	4	q=4,l=1	
2	2	1	5	0	4	0	4	q=4,l=2	
1	1	2	2						
1	2	1	4	1	0	6	12	q=5,l=1	
2	2	2	5	11	12	6	12	q=5,l=2	
2	2	1	5						
2	2	2	5						
2	2	1	5						
2	1	1	2	j=1	j=2	j=3			
2	1	2	3	1	1	1	0.1333	q=1,l=1	
2	2	2	5	0	0	0	0.1333	q=1,l=2	
2	1	1	2	0.125	1	0.875	0.2667	q=2,l=1	
2	2	2	5	0.875	0	0.125	0.2667	q=2,l=2	
2	1	1	2						
2	2	2	5	0	1	0	0.0667	q=3,l=1	
2	2	1	5	1	0	1	0.0667	q=3,l=2	
1	2	1	4						
2	1	1	2	1	0	1	0.1333	q=4,l=1	
2	2	1	5	0	1	0	0.1333	q=4,l=2	
1	1	1	1	0.0833	0	0.5	0.4	q=5,l=1	
2	2	1	5	0.9167	1	0.5	0.4	q=5,l=2	
2	1	1							

$$p(q) = \frac{H_q}{H}$$

$$p(x_j | q) = \frac{H_{qjl}}{H_q}$$

Если входные переменные являются непрерывными, то для применения данного алгоритма их следует разбить на классы. В случае равномерного распределения входных величин приведение к дискретному виду проводится по формуле:

$$x_j^{discr} = l_j = Trunc \left[(L_j - 1) \frac{x_j - x_j^{\min}}{x_j^{\max} - x_j^{\min}} \right] + 1, \quad \text{где } Trunc \text{ обозначает}$$

отбрасывание дробной части числа; x_j^{\min} , x_j^{\max} - минимальные и максимальные значения переменных; L_j - количество классов (интервалов), на которое разбивается входная переменная.

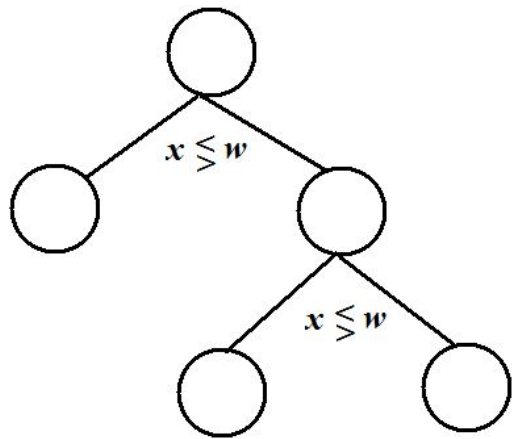
Метод классификации, основанный на деревьях решений

Деревья решений - это способ представления правил в иерархической, последовательной структуре, где каждому объекту соответствует единственный узел, дающий решение.

Под правилом понимается логическая конструкция, представленная в виде *if A then B*.

$$X = \{ \mathbf{x}^j, C_k^j \}, j = \overline{1, N}; k = \overline{1, K}$$

МНОЖЕСТВО ДАННЫХ



$x_i > (w_{ij}), i = \overline{1, m}; j = \overline{1, n}$ -условие
-разделения
по атрибуту

$$w_{ij} = x_i^{\min} + (x_i^{\max} - x_i^{\min}) \frac{j}{M_i}, i = \overline{1, m}; j = \overline{1, M_i}$$

$$P_{liq}^k, k = \overline{1, K}; i = \overline{1, m-1}; q = \overline{1, n}; l = \overline{1, 2}$$

- вероятность

принадлежности

классу k по атрибуту i

$$x_i \underset{<}{>} (w_{ij}), i = \overline{1, m-1}; j = \overline{1, n}.$$

и q -му пороговому значению

P_l^k - вероятность попадания в класс k .

Разбиению множества X по проверке t соответствует

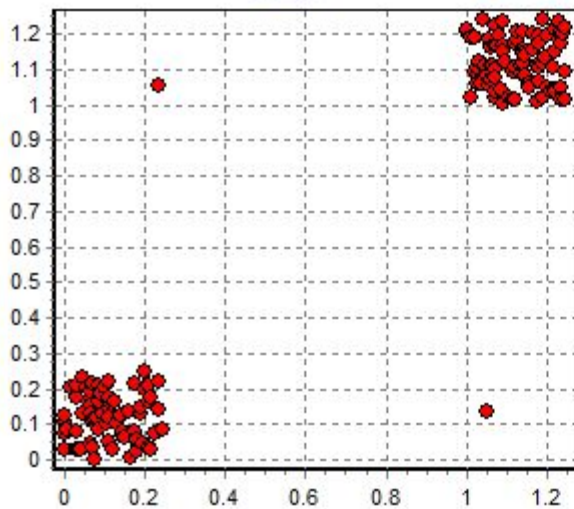
выражение для энтропии Шеннона

$$H_{iq} = - \sum_l \sum_{k=1}^K P_{liq}^k \log_2 P_{liq}^k$$

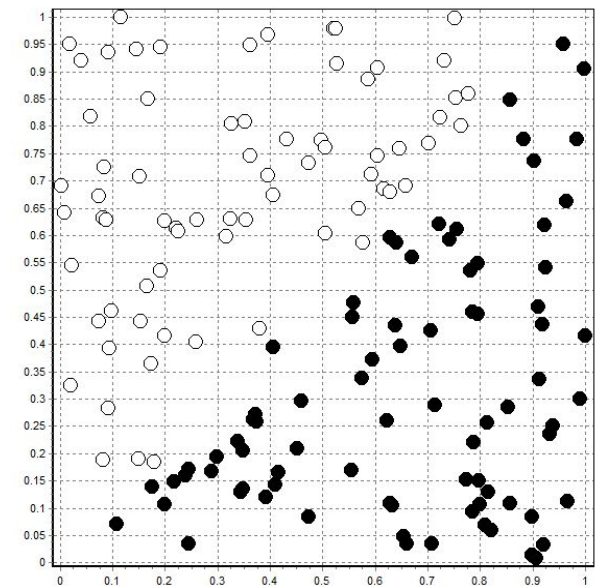
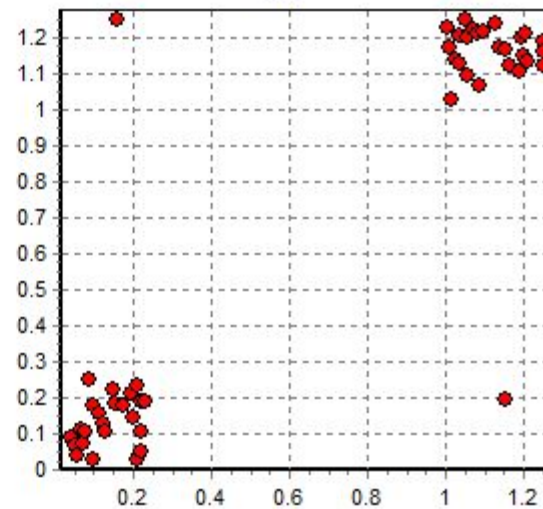
Критерий выбора

$$H_{iq} \rightarrow \min$$

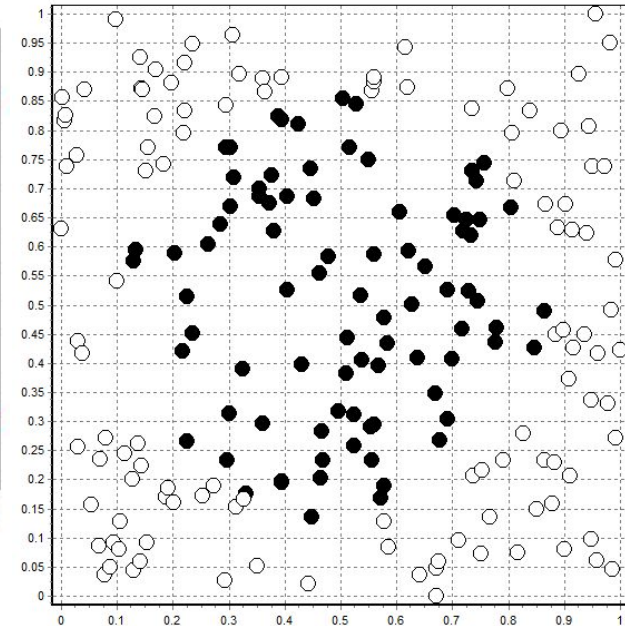
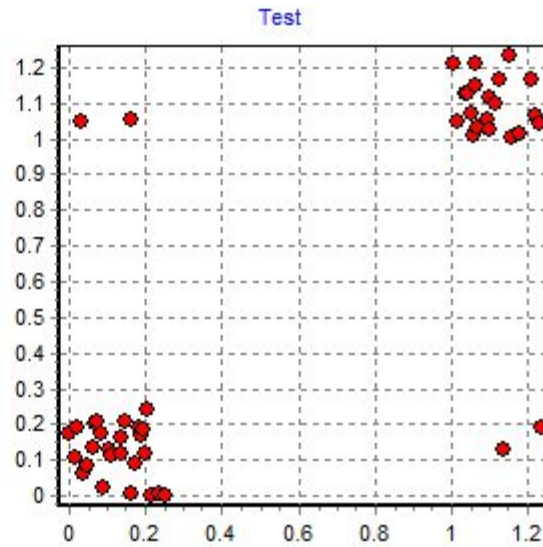
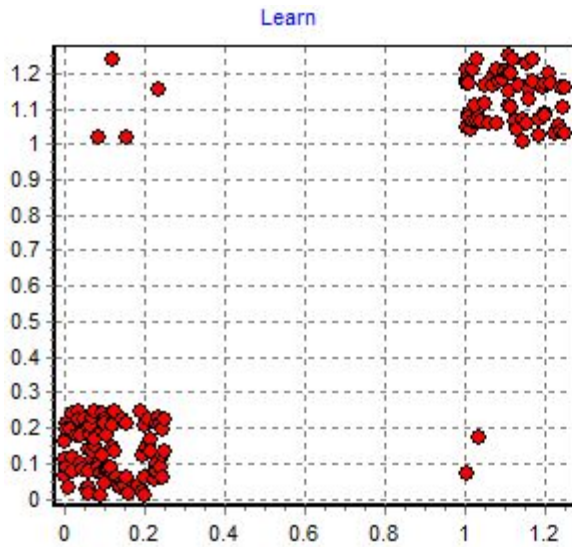
Learn



Test

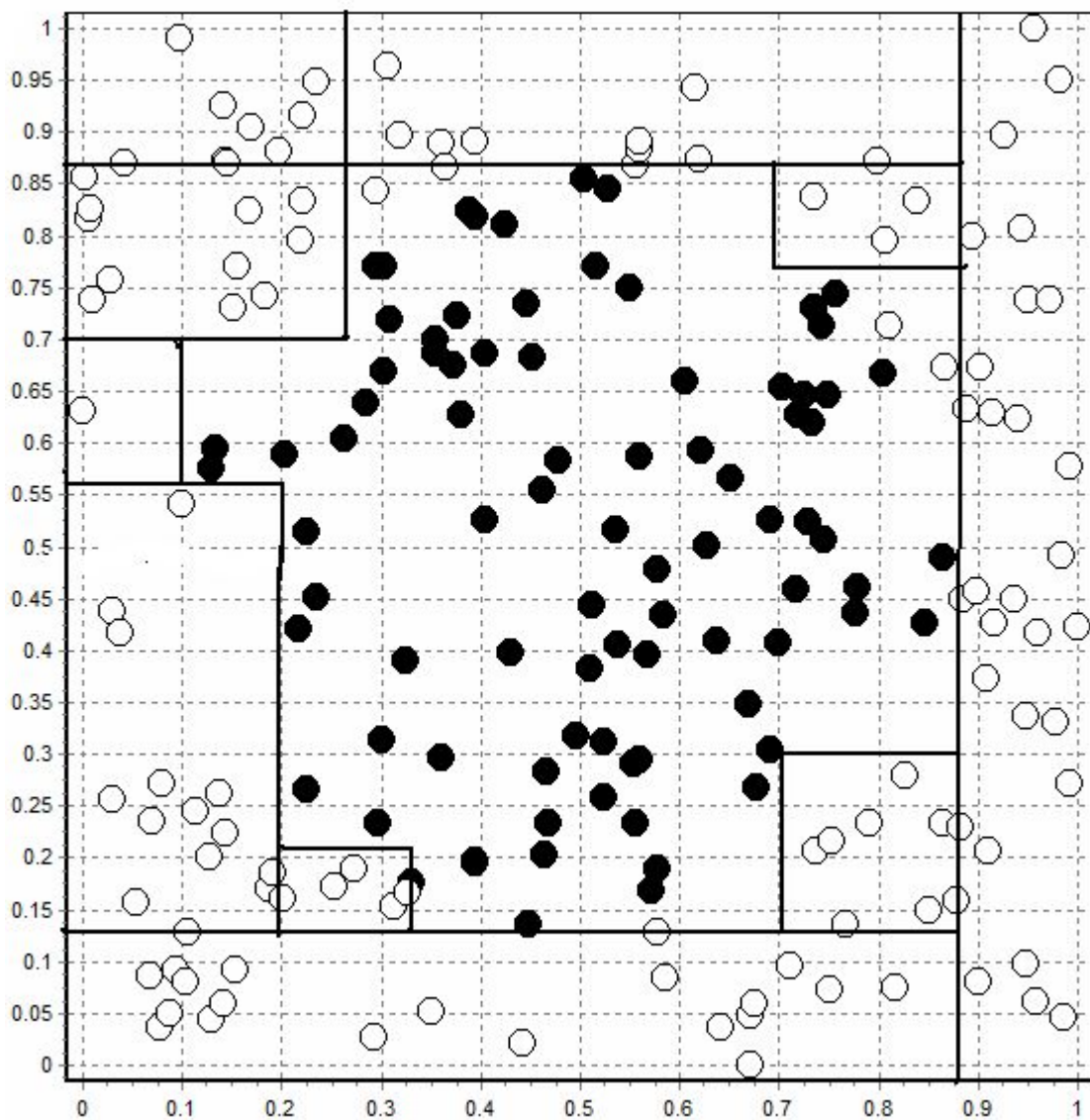


- 0 if $X[1] \geq 0.36$ AND $X[0] \geq 1.64$ then $Y = 1$
- 1 if $X[0] \geq 0.13$ AND $X[0] < 1.64$ AND $X[1] \geq 1.60$ then $Y = 0$
- 2 if $X[0] < 1.64$ AND $X[1] < 1.60$ AND $X[1] \geq 0.61$ AND $X[0] \geq 1.18$ then $Y = 1$
- 3 if $X[1] < 1.60$ AND $X[1] \geq 0.61$ AND $X[0] \geq 1.08$ AND $X[0] < 1.18$ then $Y = 0$
- 4 if $X[1] < 1.60$ AND $X[1] \geq 0.61$ AND $X[0] \geq 0.64$ AND $X[0] < 1.08$ then $Y = 0$
- 5 if $X[0] \geq 0.13$ AND $X[1] < 1.60$ AND $X[1] \geq 0.61$ AND $X[0] < 0.64$ then $Y = 0$
- 6 if $X[1] \geq 0.36$ AND $X[0] < 1.64$ AND $X[1] < 0.61$ AND $X[0] \geq 0.41$ then $Y = 1$
- 7 if $X[1] \geq 0.36$ AND $X[0] \geq 0.13$ AND $X[1] < 0.61$ AND $X[0] < 0.41$ then $Y = 0$
- 8 if $X[1] \geq 0.36$ AND $X[0] \geq -0.37$ AND $X[0] < 0.13$ then $Y = 0$
- 9 if $X[1] \geq 0.36$ AND $X[0] \geq -1.18$ AND $X[0] < -0.37$ then $Y = 0$
- 10 if $X[1] \geq 0.36$ AND $X[0] < -1.18$ then $Y = 0$
- 11 if $X[1] < 0.36$ AND $X[0] \geq 0.40$ then $Y = 1$
- 12 if $X[1] < 0.36$ AND $X[0] < 0.40$ AND $X[0] \geq -1.54$ AND $X[1] \geq -0.01$ then $Y = 0$
- 13 if $X[0] < 0.40$ AND $X[1] \geq -1.50$ AND $X[1] < -0.01$ AND $X[0] \geq -0.38$ then $Y = 1$
- 14 if $X[0] \geq -1.54$ AND $X[1] < -0.01$ AND $X[0] < -0.38$ AND $X[1] \geq -0.72$ then $Y = 0$
- 15 if $X[1] \geq -1.50$ AND $X[0] < -0.38$ AND $X[1] < -0.72$ AND $X[0] \geq -1.22$ then $Y = 1$
- 16 if $X[1] \geq -1.50$ AND $X[0] \geq -1.54$ AND $X[1] < -0.72$ AND $X[0] < -1.22$ then $Y = 0$
- 17 if $X[1] < 0.36$ AND $X[1] \geq -1.50$ AND $X[0] < -1.54$ then $Y = 0$
- 18 if $X[0] < 0.40$ AND $X[1] < -1.50$ then $Y = 1$



- 0 if $X[1] \geq 1.39$ then $Y = 0$
- 1 if $X[1] < 1.39$ AND $X[1] \geq -1.52$ AND $X[0] \geq 1.43$ then $Y = 0$
- 2 if $X[1] < 1.39$ AND $X[0] < 1.43$ AND $X[1] \geq -0.07$ AND $X[0] \geq 1.23$ then $Y = 0$
- 3 if $X[1] < 1.39$ AND $X[0] \geq 0.91$ AND $X[1] \geq -0.07$ AND $X[0] < 1.23$ then $Y = 1$
- 4 if $X[0] < 1.43$ AND $X[0] \geq 0.91$ AND $X[1] \geq -1.06$ AND $X[1] < -0.07$ then $Y = 1$
- 5 if $X[1] \geq -1.52$ AND $X[0] < 1.43$ AND $X[0] \geq 0.91$ AND $X[1] < -1.06$ then $Y = 0$
- 6 if $X[1] < 1.39$ AND $X[1] \geq -1.52$ AND $X[0] \geq 0.25$ AND $X[0] < 0.91$ then $Y = 1$
- 7 if $X[1] < 1.39$ AND $X[1] \geq -1.52$ AND $X[0] \geq 0.04$ AND $X[0] < 0.25$ then $Y = 1$
- 8 if $X[1] < 1.39$ AND $X[1] \geq -1.52$ AND $X[0] \geq -0.28$ AND $X[0] < 0.04$ then $Y = 1$
- 9 if $X[1] < 1.39$ AND $X[1] \geq -1.52$ AND $X[0] \geq -0.63$ AND $X[0] < -0.28$ then $Y = 1$
- 10 if $X[1] < 1.39$ AND $X[0] < -0.63$ AND $X[1] \geq -1.00$ AND $X[0] \geq -1.01$ then $Y = 1$
- 11 if $X[1] < 1.39$ AND $X[0] \geq -1.47$ AND $X[0] < -1.01$ AND $X[1] \geq 0.67$ then $Y = 0$
- 12 if $X[0] \geq -1.47$ AND $X[1] \geq -1.00$ AND $X[0] < -1.01$ AND $X[1] < 0.67$ then $Y = 1$
- 13 if $X[1] \geq -1.52$ AND $X[0] \geq -1.47$ AND $X[0] < -0.63$ AND $X[1] < -1.00$ then $Y = 0$
- 14 if $X[1] < 1.39$ AND $X[1] \geq -1.52$ AND $X[0] < -1.47$ then $Y = 0$
- 15 if $X[1] < -1.52$ then $Y = 0$

Разделение области



Дерево решений

