

Интервальный метод классификации состояний в задачах медицинской диагностики

А. С. Подружко

Аннотация. В работе представлен интервальный подход к решению задачи классификации по нескольким показателям применительно к проблеме медицинской диагностики. В отличие от вероятностных методов здесь используется вариационный подход, основанный на выпуклой аппроксимации поля данных. Классификация сводится к решению пары задач линейной оптимизации при соблюдении необходимых условий полноты охвата статистики и выпуклости оболочки ограничений. Определены правила оценки информативности признаков и формирования заключений по нескольким признакам. Приведены результаты решений задач для случая 3 показателей.

Ключевые слова: интервальная классификация, оптимизация, информативность признака, решающие правила.

Введение

Решение задач классификации состояний обычно строится на основе вероятностных представлений с использованием функций правдоподобия различного вида или иных методов детерминационного анализа [1–3]. При этом само решение предполагает, как правило, выбор одной из двух гипотез и вероятностные оценки каждой из них. Полученные результаты справедливы, строго говоря, только для множества будущих решений. Однако часто такого рода правила применяются к отдельным объектам или событиям, в частности, в медицинской диагностике при формировании заключений по каждому конкретному пациенту. Полученные заключения в этом случае вызывают определенные сомнения.

В данной работе представлен качественно иной, интервальный подход к решению задачи классификации, где решающее правило выбора строится без использования каких-либо вероятностных характеристик на основе неявных регрессий.

1. Формулировка задачи интервальной классификации

Изначально формулировку задачи классификации состояний можно представить следующим образом. Будем полагать, что анализируемый объект (клетка) может находиться в двух основных состояниях S : A — здоровая клетка и B — больная клетка.

Распознавание состояния осуществляется на основе семейства признаков

$$P = \{p_1, p_2, \dots, p_m, \dots, p_M\}$$

и набора прецедентов (статистики) по каждому из состояний Ω_A и Ω_B объемом N_A и N_B соответственно:

$$\Omega_A = \{P_1^{(A)}, P_2^{(A)}, \dots, P_i^{(A)}, \dots, P_{N_A}^{(A)}\}$$

и

$$\Omega_B = \{P_1^{(B)}, P_2^{(B)}, \dots, P_j^{(B)}, \dots, P_{N_B}^{(B)}\}.$$

Также будем полагать, что по отдельным признакам (при прочих равных условиях) существует явное отношение предпочтения, которое может быть выражено следующим образом:

- чем больше значение любого признака p_m , тем более предпочтительно состояние A ,
- чем меньше его значение, тем предпочтительней состояние B .

Если же имеет место обратное, то достаточно изменить знак значения признака. Эта гипотеза находит практическое подтверждение во многих случаях и, в частности, в медицинской диагностике.

В рамках процедуры интервальной классификации помимо основных состояний клетки A и B допускаются также неопределенные ситуации N , которые можно интерпретировать как третье промежуточное состояние, т. е. $S = \{A, B, N\}$. При этом целью интервальной классификации является максимальное сужение области неопределенности в поле признаков за счет отбора среди них наиболее информативных и последовательного расширения их

состава и, в конечном счете, построения эффективных правил распознавания состояний клетки.

2. Оценка информативности признаков на расширенном поле данных

На заданном поле данных $\Omega_A \cup \Omega_B$ определим по каждому из показателей p_m их минимальные и максимальные значения по двум видам прецедентов:

$$\begin{aligned}
 A: \quad & a_m^{(1)} = \min(p_m^{(A)})_i, \\
 & a_m^{(2)} = \max(p_m^{(A)})_i, \\
 & i = 1, 2, \dots, N_A; \\
 B: \quad & b_m^{(1)} = \min(p_m^{(B)})_j, \\
 & b_m^{(2)} = \max(p_m^{(B)})_j, \\
 & j = 1, 2, \dots, N_B.
 \end{aligned} \tag{1}$$

Доопределим исходное поле данных по каждому из показателей до интервала в пределах найденных границ $I_m^{(A)} = [a_m^{(1)}, a_m^{(2)}]$, $I_m^{(B)} = [b_m^{(1)}, b_m^{(2)}]$ и будем рассматривать интервалы как аппроксимацию поля данных. Совмещение этих интервалов на оси измерения показателя p_m позволяет (в большинстве случаев) выделить три области: $D_m^{(A)} = [b_m^{(2)}, a_m^{(2)}]$ — область, относящаяся к состоянию А, $D_m^{(B)} = [b_m^{(1)}, a_m^{(1)}]$ — к состоянию В, $D_m^{(N)} = [a_m^{(1)}, b_m^{(2)}]$ — область неопределенности, а кроме того, указать область определения правил классификации $D_m^{(O)} = [b_m^{(1)}, a_m^{(2)}]$ по этому показателю. Соответственно, появляется необходимость добавить в общий список состояний еще одно О – состояние, случай, когда показатели анализируемой ситуации находятся за пределами области определения $D_m^{(O)}$ и применение правил классификации невозможно, т. е. $S = \{A, B, N, O\}$.

Представленная схема разбиения поля данных существует при условии, когда пересечение интервалов по прецедентам $I_m^{(A)}$ и $I_m^{(B)}$ будет не пустым: $I_m^{(A)} \cap I_m^{(B)} = D_m^{(N)} \neq \emptyset$, а сами интервалы не поглощают друг друга, т. е. $I_m^{(A)} \not\subset I_m^{(B)}$ и $I_m^{(B)} \not\subset I_m^{(A)}$. В противном случае — признак будем считать не информативным. При соблюдении необходимых условий схема разбиения имеет оговоренную структуру

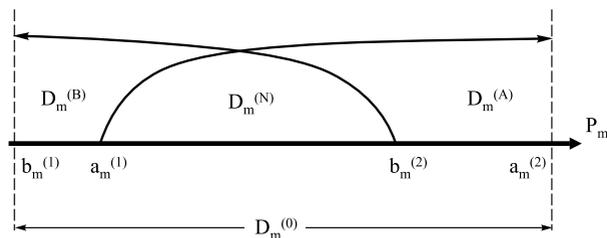


Рис. 1. Схема пересечений проекций показателей

и может быть использована в первую очередь для предварительной оценки информативности показателей

В качестве меры информативности k_m будем рассматривать отношение ширины интервала неопределенности $D_m^{(N)}$ к величине интервала области определения $D_m^{(O)}$:

$$k_m = \frac{b_m^{(2)} - a_m^{(1)}}{a_m^{(2)} - b_m^{(1)}}, \quad m = 1, 2, \dots, M. \tag{2}$$

Чем меньше ширина интервала неопределенности, тем меньше значение k_m и тем выше информативность показателя при построении решающих правил. И наоборот, чем больше значение k_m , тем больше неопределенность при определении состояния S и тем меньше информативность показателя. Используя эту оценку информативности можно ранжировать показатели по информативности и тем самым упорядочить их использование при классификации.

Помимо упорядочивания практическое использование интервального представления поля данных по отдельным признакам целесообразно только для проверки анализируемых ситуаций на возможность классификации (О – состояние). Распознавание ситуаций на множестве показателей, с использованием интервалов на многомерном кубе представляется заведомо бесперспективным.

3. Общее описание интервальной процедуры классификации состояний

Для того, чтобы учесть неявные связи между показателями при распознавании ситуаций воспользуемся обобщенным интервальным методом (DEA — технологией [4–6]), основанным на понятии фронта экстремальных значений. Идея метода заключается в построении на M -мерном поле данных двух границ (фронт) Γ_A и Γ_B в виде выпуклых многогранников, охватывающих все точки статистики по каждой группе прецедентов А и В (принцип полноты охвата статистики). Экстремальное свойство границ состоит в том, что они отделяют на поле данных области минимального объема, соответствующие А и В прецедентам соответственно. Пересечение этих границ и областей по прецедентам является основой для формирования решающих правил классификации.

Рассмотрим вначале двухмерную задачу классификации $\mathbf{P} = \{p_1, p_2\}$, наглядно демонстрирующую особенности данного метода. На заданной статистике Ω_A и Ω_B экстремальное положение границ Γ_A и Γ_B можно определить как решение следующей

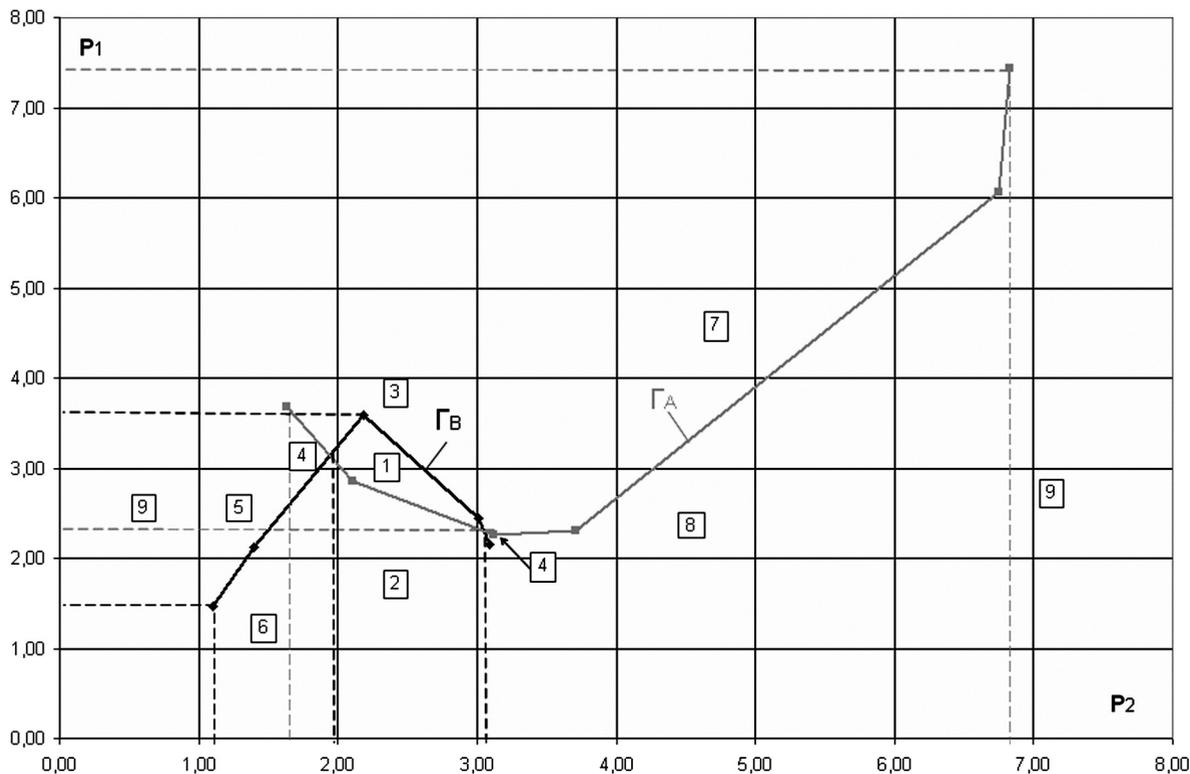


Рис. 2. Диаграмма состояний

пары оптимизационных задач с учетом условий выпуклости оболочки:

$$\begin{aligned}
 & \Gamma_A & \Gamma_B \\
 & \min p_1^{(A)}(p_2^{(A)}), & \max p_1^{(B)}(p_2^{(B)}), \\
 & \sum_{i=1}^{N_A} \lambda_i(p_1^{(A)})_i = p_1^{(A)}, & \sum_{j=1}^{N_B} \mu_j(p_1^{(B)})_j = p_1^{(B)}, \\
 & \sum_{i=1}^{N_A} \lambda_i(p_2^{(A)})_i = p_2^{(A)}, & \sum_{j=1}^{N_B} \mu_j(p_2^{(B)})_j = p_2^{(B)}, \\
 & \sum_{i=1}^{N_A} \lambda_i = 1, \lambda_i \geq 0, & \sum_{j=1}^{N_B} \mu_j = 1, \mu_j \geq 0, \\
 & i = 1, 2, \dots, N_A, & j = 1, 2, \dots, N_B, \\
 & p_2^{(A)} = p_2^*; p_2^* \in I_2^{(A)}; & p_2^{(B)} = p_2^{**}, p_2^{**} \in I_2^{(B)}.
 \end{aligned} \quad (3)$$

Результатом решения будут два кусочно-линейных фронта $\Gamma_A: p_1^{(A)}(p_2^{(A)})$ и $\Gamma_B: p_1^{(B)}(p_2^{(B)})$, образующих на плоскости (p_1, p_2) диаграмму состояний (рис. 2).

На диаграмме нетрудно выделить ряд областей, соответствующих различным сочетаниям пар (p_1, p_2) и сформулировать логическое правило классификации состояния S_X с заданными значениями показателей $(p_1^{(X)}, p_2^{(X)})$. При этом будем использовать ого-

воренные правила предпочтения: если точка на диаграмме будет находиться выше фронта Γ_A , то предпочтительней ее отнести к состоянию А, если же ниже — то к обратному состоянию $\neg A$. Для другого фронта Γ_B действует обратное правило: если точка ниже фронта, то предпочтительней отнести ее к состоянию В, если же выше, то к обратному состоянию $\neg B$. Также учтены ситуации, когда анализируемое значение показателей выходит за пределы области определения границ (О-состояния). Для произвольной пары (p_1, p_2) правило классификации теперь можно представить в виде табл. 1.

Таблица 1

| n | P1 | P2 | S |
|---|----------|----------|---|
| 1 | A | B | N |
| 2 | $\neg A$ | B | B |
| 3 | A | $\neg B$ | A |
| 4 | $\neg A$ | $\neg B$ | N |
| 5 | O | $\neg B$ | A |
| 6 | O | B | B |
| 7 | A | O | A |
| 8 | $\neg A$ | O | B |
| 9 | O | O | O |

Как следует из диаграммы и табл. 1 в общем случае можно выделить 9 непересекающихся областей и столько же условий, определяющих диагностируемое состояние. Логика этих правил достаточно очевидна, хотя и не однозначна. В ряде случаев они могут быть видоизменены в зависимости от особенностей стратегии диагностики и лечения. Однако в любом случае решение будет определяться парой оценок состояния. В практическом плане процедура классификации сводится к проверке заданной пары показателей $(p_1^{(X)}, p_2^{(X)})$ на полученной диаграмме состояний и последующего использования решающего правила на основе табл. 1.

При большом объеме исходной статистики определенные проблемы могут возникнуть в связи с высокой размерностью оптимизационных моделей (3). Однако здесь всегда удастся выделить достаточные статистики по каждому фронту $G_A \subset \Omega_A$ и $G_B \subset \Omega_B$, которые в полной мере определяют форму этих фронтов Γ_A и Γ_B . Объем точек, входящих в достаточные статистики, как правило, невелик. В рассмотренном примере (рис. 1) G_A содержит 6 точек, а G_B — 5. Это позволяет на этапе пополнения статистической базы отбирать на каждом шаге только необходимые точки G_A и G_B , а затем, при повторном решении задач (3), дополнять достаточные статистики новыми данными с последующей редукцией состава достаточных статистик. Таким образом, за счет многошаговой процедуры удается решить проблему размерности при определении вида диаграммы состояний.

Рассмотренную процедуру классификации нетрудно обобщить на случай трех и более показателей. Однако следует заметить, что в большинстве случаев на практике (не обязательно в медицинской диагностике) решения базируются на 3–4 наиболее информативных показателях. Выбор этих показателей может осуществляться как опытным путем, так и с использованием аналитических методов, например, с помощью представленной здесь схемы ранжирования показателей по информативности.

Геометрически представить границы Γ_A и Γ_B в трехмерном пространстве (p_1, p_2, p_3) затруднительно. Однако ряд общих свойств этих фронтов можно проанализировать по их проекциям на плоскости (p_1, p_2) при фиксированных значениях третьего показателя p_3 . Эти проекции можно получить как решение оптимизационных задач, аналогичных (3), при следующих условных критериях:

$$\Gamma_A : \min p_1^{(A)}(p_2^{(A)} | p_3^{(A)}) \quad \Gamma_B : \max p_1^{(B)}(p_2^{(B)} | p_3^{(B)}), \quad (4)$$

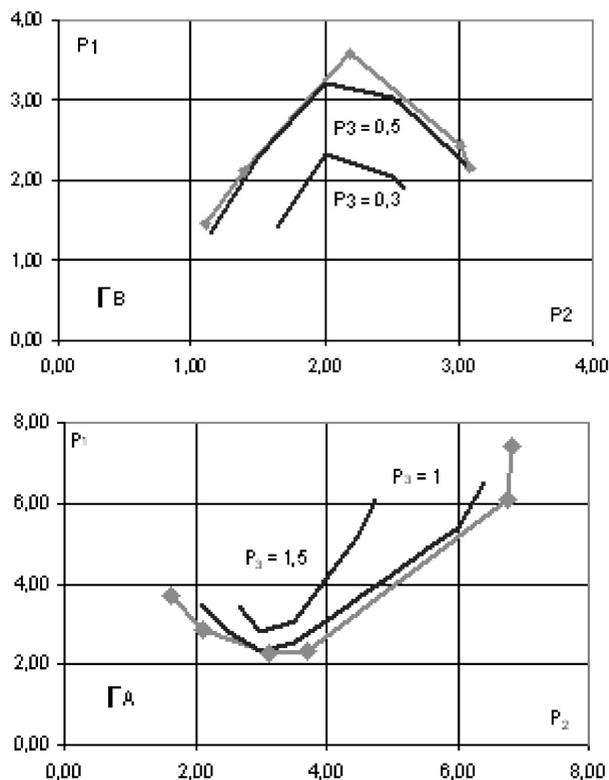


Рис. 3. Проекция границ разных уровней

и дополнительных ограничениях по третьему показателю $p_3^{(A)}$ и $p_3^{(B)}$:

$$\sum_{i=1}^{N_A} \lambda_i(p_3^{(A)})_i = p_3^{(A)},$$

$$\sum_{i=1}^{N_B} \lambda_i(p_3^{(B)})_i = p_3^{(B)}, \quad (5)$$

$$p_3^{(A)} \in I_3^{(A)},$$

$$p_3^{(B)} \in I_3^{(B)}.$$

Проекция представляет собой вложения в огибающую, полученную для двухмерного случая (рис. 3). Соответственно, привлечение для классификации третьего показателя позволяет уменьшить область пересечения фронтов и неопределенность в принятии решения. Однако при этом может существенно сужаться область определения по каждой из границ, что также необходимо учитывать при построении решающих правил.

Таким образом, определение состояния S_X с показателями $(p_1^{(X)}, p_2^{(X)}, p_3^{(X)})$ в конечном счете сводится к вычислению значения координат точек a_1 и b_1 на проекциях фронтов Γ_A и Γ_B при заданных значениях двух других координат $(p_2^{(X)}, p_3^{(X)})$ и последующему сопоставлению с фактическим значе-

нием первого показателя $p_1^{(X)}$. Если $p_1^{(X)} \geq a_1$, то первой из гипотез принимается решение А, в противном случае обратное $\neg A$. Если $p_1^{(X)} \leq b_1$, то по второй гипотезе принимается решение В, а в противном случае $\neg B$. Если при заданных показателях $(p_2^{(X)}, p_3^{(X)})$ задачи (3), (4) (одна или обе) окажутся несовместны, то соответствующей из гипотез присваивается О-статус. Это означает, что пара показателей $(p_2^{(X)}, p_3^{(X)})$ находится за пределами области определения фронта. Теперь анализируемое состояние S_X можно классифицировать, пользуясь табл. 1.

Последующее расширение состава показателей ведет лишь к увеличению размерности основной вычислительной задачи (3) и не влияет на процедуру формирования решений. Выбор конкретного числа этих показателей, особенно в медицинских приложениях, должен устанавливаться экспериментально в зависимости от характера диагностической процедуры.

4. Результаты тестирования процедуры интервальной классификации

Эффективность представленной здесь процедуры анализировалась на основе реальных данных, полученных при диагностике заболеваний крови (лейкозов). Были отобраны по критерию информативности три признака и, соответственно, три группы прецедентов А и В вида (по 70 прецедентов каждого вида), которые в дальнейшем использовались при построении классификатора (Приложение 1). Для тестирования использовались выборки по 30 точек выявленных прецедентов каждого вида, естественно, не используемых в классификаторе (Приложения 2 и 3).

Полученные результаты оказались весьма обнадеживающими. В частности, при классификации только по одному наиболее информативному признаку p_1 надежность распознавания больных клеток (В – клеток) составила 73,3 % при ошибке 3,3 % и 23,3 % неопределенных ситуаций. Это во многом объясняется хорошим значением показателя информативности $k_1 = 0,156$ по первому (в ранжированном ряде) признаку p_1 при достаточно широкой области определения по этому признаку $D_1^{(O)} = [0,32, 8,78]$ и относительно небольшом интервале неопределенности $D_1^{(N)} = [2,26, 3,58]$. В то же время распознавание здоровых клеток (А – клеток) оказалось существенно хуже. Надежность решений составила только 33,3 % при ошибке 6,7 % и 60,0 % неопределенных ситуаций.

Показатели информативности по двум другим признакам p_2 и p_3 существенно хуже, $k_2 = 0,254$ и $k_3 = 0,324$. Тем не менее, классификация на поле двух признаков p_1 и p_2 позволила заметно повысить надежность распознавания, особенно А-клеток до 73,3 % при ошибке 16,7 % и 10,0 % неопределенных ситуаций. Для В-клеток показатель надежности оказался еще выше и составил 90,0 % при ошибке 3,3 % и 6,7 % неопределенных ситуаций. На следующем шаге, на поле трех признаков p_1 , p_2 и p_3 удалось полностью устранить неопределенность в классификации и получить надежность по А-клеткам 83,3 % при ошибке 16,7 % и В-клеткам 93,3 % при ошибке 6,7 %. Следует заметить, что в рассмотренном примере ошибки классификации, в частности В-клеток, связаны с явно аномальными значениями признаков в анализируемой группе прецедентов. Скорее всего, это погрешности измерений, что является предметом отдельного исследования. В целом же эффективность процедуры интервальной классификации на данной статистике оказалась весьма высокой, особенно при распознавании больных клеток, что наиболее существенно с точки зрения стратегии лечения.

Следует также отметить, что при увеличении размерности поля признаков имеет место наследование определенных состояний клеток. Если состояние А или В клетки выявлено на поле меньшей размерности, то оно же сохранится и на поле более высокой размерности. Это объясняется общими свойствами проектирования метрических пространств и может быть использовано для упрощения процедуры классификации.

Заключение

Представленная здесь процедура интервальной классификации может быть использована не только в медицине, но и в любых иных приложениях, где решающие правила строятся на поле прецедентов двух видов. В частности, при прогнозировании финансовой состоятельности компаний по обобщенным показателям их деятельности.

Литература

1. Пугачев В. С. Теория вероятностей и математическая статистика. М.: Наука, 1979.
2. Розанов Ю. А. Теория вероятностей, случайные процессы и математическая статистика. М.: Наука, 1985.
3. Босс В. Лекции по математике. Т. 4: Вероятность, информация, статистика. М.: ЛКИ/URSS, 2008.
4. Banker R. D., Charnes A., Cooper W. W. Some models for estimating technical and scale efficiency in data envelopment analysis (DEA) // Management Science. 1984. V. 30. № 9.

5. Кривоножко В. Е., Пропой А. И. и др. Анализ эффективности функционирования сложных систем // Автоматизация проектирования. № 1. 1999. С. 2–27.
6. Подружко А. А., Подружко А. С. Интервальное представление надежных регрессий // Динамика неоднородных систем. Вып. 10. М.: КомКнига/URSS, 2006.

Подружко Александр Степанович. С. н. с. ИСА РАН, к. т. н. Окончил МИФИ в 1969 г. Количество печатных работ: более 60, 1 монография. Область научных интересов: оптимизационное моделирование, интервальный анализ, вариационные методы статистики, риск-менеджмент.
E-mail: ASP45@aknet.ru

Приложение 1

Исходные данные для расчетов по двум видам прецедентов

| А | | | В | | |
|------|------|------|------|------|------|
| 1 | 2 | 3 | 1 | 2 | 3 |
| 2,74 | 3,34 | 0,99 | 4,61 | 4,59 | 0,47 |
| 3,32 | 2,66 | 1,48 | 3,68 | 1,64 | 0,62 |
| 3,00 | 3,05 | 0,92 | 5,84 | 3,97 | 0,48 |
| 5,02 | 3,14 | 1,74 | 5,90 | 4,05 | 0,65 |
| 2,26 | 3,12 | 1,08 | 5,60 | 3,91 | 0,48 |
| 3,00 | 2,81 | 1,03 | 5,50 | 3,27 | 0,54 |
| 3,13 | 2,68 | 1,18 | 5,56 | 3,61 | 0,70 |
| 2,30 | 3,06 | 0,88 | 8,28 | 5,29 | 0,46 |
| 2,30 | 3,71 | 1,26 | 4,83 | 2,44 | 0,74 |
| 2,37 | 3,20 | 1,49 | 8,57 | 4,80 | 0,77 |
| 3,99 | 3,98 | 1,07 | 5,09 | 4,76 | 0,78 |
| 6,12 | 2,47 | 0,87 | 7,49 | 3,72 | 1,34 |
| 3,63 | 4,29 | 0,84 | 5,46 | 3,28 | 0,79 |
| 3,77 | 3,59 | 1,26 | 7,51 | 5,41 | 0,90 |
| 7,79 | 4,35 | 1,13 | 7,38 | 3,61 | 0,61 |
| 5,52 | 2,56 | 0,80 | 6,92 | 4,38 | 0,79 |
| 6,18 | 3,93 | 0,83 | 5,74 | 3,03 | 0,82 |
| 4,39 | 3,74 | 0,78 | 4,65 | 2,91 | 0,96 |
| 2,85 | 2,11 | 0,60 | 8,68 | 4,60 | 0,90 |
| 4,77 | 3,09 | 0,51 | 7,43 | 6,83 | 1,19 |
| 3,41 | 2,79 | 0,53 | 4,77 | 2,63 | 0,86 |
| 4,50 | 3,49 | 0,60 | 6,06 | 6,75 | 1,15 |
| 2,81 | 2,85 | 0,49 | 4,59 | 2,88 | 1,25 |
| 4,05 | 2,49 | 0,71 | 5,85 | 3,71 | 1,12 |
| 2,74 | 2,48 | 0,47 | 4,25 | 3,49 | 0,73 |
| 4,31 | 2,64 | 0,66 | 4,88 | 4,73 | 0,90 |
| 7,23 | 4,15 | 0,60 | 4,84 | 5,26 | 1,09 |
| 6,09 | 4,17 | 0,71 | 3,95 | 2,45 | 0,96 |
| 4,39 | 3,43 | 0,65 | 5,42 | 3,84 | 0,93 |
| 4,93 | 3,77 | 0,49 | 4,78 | 2,98 | 0,64 |
| 5,90 | 2,92 | 0,84 | 4,28 | 2,22 | 0,78 |
| 4,18 | 2,91 | 0,66 | 7,19 | 4,84 | 1,22 |
| 4,93 | 3,45 | 0,56 | 8,78 | 4,12 | 1,05 |
| 4,40 | 3,45 | 0,83 | 3,60 | 2,77 | 1,18 |
| 3,88 | 3,56 | 0,70 | 7,72 | 4,07 | 0,84 |
| 1,77 | 2,08 | 0,59 | 0,86 | 1,75 | 0,51 |
| 0,66 | 1,47 | 0,82 | 0,88 | 1,87 | 0,40 |
| 2,45 | 3,01 | 0,57 | 1,58 | 1,67 | 0,41 |
| 1,84 | 2,34 | 0,63 | 2,86 | 1,85 | 0,35 |
| 2,60 | 1,93 | 0,65 | 1,66 | 1,51 | 0,43 |
| 2,89 | 1,85 | 0,94 | 1,63 | 2,08 | 0,34 |
| 3,58 | 2,18 | 0,58 | 1,98 | 1,84 | 0,40 |
| 0,98 | 1,53 | 0,89 | 1,03 | 1,20 | 0,47 |
| 0,91 | 1,38 | 0,61 | 3,10 | 2,06 | 0,32 |
| 1,85 | 1,75 | 0,56 | 1,41 | 2,45 | 0,34 |
| 2,13 | 2,45 | 0,69 | 1,93 | 1,51 | 0,33 |
| 1,83 | 1,84 | 0,83 | 0,70 | 1,21 | 0,65 |
| 0,85 | 1,51 | 0,78 | 1,38 | 1,76 | 0,29 |
| 2,67 | 1,98 | 0,41 | 1,24 | 2,16 | 0,43 |
| 1,15 | 1,17 | 0,68 | 1,90 | 1,69 | 0,38 |
| 2,01 | 2,51 | 0,71 | 1,20 | 1,75 | 0,31 |
| 2,05 | 3,08 | 0,36 | 2,36 | 1,82 | 0,43 |
| 0,32 | 1,53 | 0,72 | 1,15 | 1,93 | 0,42 |
| 0,50 | 1,27 | 0,49 | 1,46 | 1,10 | 0,34 |
| 0,95 | 2,28 | 0,65 | 2,02 | 2,24 | 0,44 |
| 0,96 | 2,00 | 0,34 | 1,48 | 1,81 | 0,40 |
| 0,49 | 1,22 | 0,67 | 0,87 | 1,44 | 0,66 |
| 0,77 | 1,24 | 0,52 | 1,32 | 2,08 | 0,36 |
| 2,15 | 3,09 | 0,56 | 1,49 | 2,71 | 0,48 |
| 0,99 | 1,94 | 0,49 | 1,30 | 1,82 | 0,48 |
| 0,97 | 2,44 | 0,49 | 2,12 | 1,40 | 0,51 |
| 1,09 | 1,80 | 0,44 | 1,46 | 1,75 | 0,39 |
| 0,91 | 2,84 | 0,48 | 0,90 | 1,55 | 0,37 |
| 0,81 | 1,94 | 0,49 | 1,76 | 1,65 | 0,38 |
| 0,48 | 1,67 | 0,48 | 1,26 | 1,92 | 0,37 |
| 0,76 | 2,23 | 0,36 | 1,00 | 2,33 | 0,60 |
| 1,79 | 2,27 | 0,26 | 1,84 | 2,26 | 0,46 |
| 0,47 | 1,71 | 0,91 | 1,89 | 2,00 | 0,54 |
| 0,90 | 2,36 | 0,40 | 0,95 | 1,42 | 0,33 |
| 1,42 | 1,60 | 0,44 | 2,41 | 2,97 | 0,79 |

Оценки информативности

| k_1 | k_2 | k_3 |
|-------|-------|-------|
| 0,156 | 0,254 | 0,324 |

Приложение 2

Результаты тестирования по состоянию А

| | Test A | | | p1 | p1+p2 | | | p1+p2+p3 | | |
|----|--------|------|------|----|-------|-------|----|----------|-------|----|
| | 4 | 6 | 1 | | S 4 | ГА | ГВ | Sx | ГА | ГВ |
| 1 | 5,24 | 4,61 | 1,00 | A | A | B | N | A | O | A |
| 2 | 5,45 | 2,98 | 1,58 | A | A | Not B | A | | | A |
| 3 | 1,97 | 2,84 | 0,78 | B | Not A | B | B | Not A | B | B |
| 4 | 3,12 | 3,72 | 0,79 | N | A | Not B | A | | | A |
| 5 | 2,47 | 2,81 | 0,70 | N | A | B | N | A | Not B | A |
| 6 | 2,53 | 2,39 | 0,62 | N | Not A | B | B | Not A | B | B |
| 7 | 3,21 | 3,06 | 0,68 | N | A | Not B | A | | | A |
| 8 | 2,41 | 3,14 | 0,64 | N | A | O | A | | | A |
| 9 | 3,82 | 4,02 | 0,59 | A | A | O | A | | | A |
| 10 | 2,95 | 1,34 | 0,78 | N | O | Not B | A | | | A |
| 11 | 2,17 | 2,03 | 0,96 | N | Not A | B | B | Not A | B | B |
| 12 | 7,90 | 3,06 | 0,66 | A | A | Not B | A | | | A |
| 13 | 3,51 | 3,73 | 0,79 | N | A | O | A | | | A |
| 14 | 3,73 | 3,06 | 0,94 | N | A | Not B | A | | | A |
| 15 | 2,67 | 3,23 | 0,48 | N | A | O | A | | | A |
| 16 | 3,57 | 2,32 | 0,71 | N | A | Not B | A | | | A |
| 17 | 3,66 | 3,28 | 0,66 | A | A | O | A | | | A |
| 18 | 1,28 | 2,29 | 0,74 | B | Not A | O | B | B | Not A | B |
| 19 | 3,17 | 2,64 | 0,63 | N | A | Not B | A | | | A |
| 20 | 4,33 | 3,74 | 0,92 | A | A | O | A | | | A |
| 21 | 5,72 | 3,99 | 0,42 | A | A | O | A | | | A |
| 22 | 3,32 | 3,38 | 0,72 | N | A | O | A | | | A |
| 23 | 2,82 | 2,88 | 0,78 | N | A | Not B | A | | | A |
| 24 | 3,39 | 2,60 | 0,78 | N | A | Not B | A | | | A |
| 25 | 2,87 | 4,04 | 0,56 | N | A | O | A | | | A |
| 26 | 5,20 | 3,13 | 0,98 | A | A | O | A | | | A |
| 27 | 4,07 | 2,16 | 0,62 | A | A | Not B | A | | | A |
| 28 | 2,80 | 2,65 | 0,80 | N | A | B | N | A | Not B | A |
| 29 | 3,62 | 4,38 | 0,67 | A | A | O | A | | | A |
| 30 | 2,61 | 2,45 | 0,54 | N | Not A | B | B | Not A | B | B |

| A | B | N |
|-------|------|-------|
| 10 | 2 | 18 |
| 33,3% | 6,7% | 60,0% |

| A | B | N |
|-------|-------|-------|
| 22 | 5 | 3 |
| 73,3% | 16,7% | 10,0% |

| A | B | N |
|-------|-------|------|
| 25 | 5 | 0 |
| 83,3% | 16,7% | 0,0% |

Приложение 3

Результаты тестирования по состоянию В

| | Test B | | | p1 | p1+p2 | | | p1+p2+p3 | | |
|----|--------|------|------|----|-------|-------|----|----------|-------|----|
| | 4 | 6 | 1 | | S 4 | ГA | ГB | Sx | ГA | ГB |
| 1 | 1,51 | 3,08 | 0,55 | B | Not A | B | B | | | B |
| 2 | 0,84 | 1,94 | 0,54 | B | Not A | B | B | | | B |
| 3 | 1,07 | 1,84 | 0,45 | B | Not A | B | B | | | B |
| 4 | 1,86 | 1,66 | 0,46 | B | Not A | B | B | | | B |
| 5 | 1,38 | 2,43 | 0,45 | B | Not A | B | B | | | B |
| 6 | 1,60 | 2,22 | 0,66 | B | Not A | B | B | | | B |
| 7 | 0,93 | 1,93 | 0,43 | B | Not A | B | B | | | B |
| 8 | 0,40 | 1,81 | 0,61 | B | Not A | B | B | | | B |
| 9 | 4,39 | 3,27 | 0,54 | A | A | O | A | A | O | A |
| 10 | 1,86 | 1,52 | 0,60 | B | O | B | B | | | B |
| 11 | 2,06 | 1,67 | 0,32 | B | Not A | B | B | | | B |
| 12 | 1,39 | 2,27 | 0,34 | B | Not A | B | B | | | B |
| 13 | 0,92 | 1,36 | 0,29 | B | O | B | B | | | B |
| 14 | 1,16 | 1,49 | 0,36 | B | O | B | B | | | B |
| 15 | 1,20 | 1,62 | 0,28 | B | O | B | B | | | B |
| 16 | 2,41 | 4,10 | 0,63 | N | Not A | O | B | | | B |
| 17 | 1,20 | 1,25 | 0,89 | B | O | B | B | | | B |
| 18 | 0,80 | 2,22 | 0,97 | B | Not A | B | B | | | B |
| 19 | 2,13 | 2,99 | 0,58 | B | Not A | B | B | | | B |
| 20 | 1,76 | 4,26 | 0,78 | B | Not A | B | B | | | B |
| 21 | 2,94 | 2,83 | 0,69 | N | Not A | Not B | N | A | Not B | A |
| 22 | 3,35 | 1,75 | 0,67 | N | Not A | Not B | N | Not A | B | B |
| 23 | 1,74 | 4,23 | 0,61 | B | Not A | O | B | | | B |
| 24 | 1,12 | 2,05 | 0,36 | B | Not A | B | B | | | B |
| 25 | 2,43 | 3,34 | 0,76 | N | Not A | O | B | | | B |
| 26 | 1,14 | 1,81 | 0,71 | B | Not A | B | B | | | B |
| 27 | 2,41 | 3,38 | 0,59 | N | Not A | O | B | | | B |
| 28 | 3,00 | 3,99 | 0,65 | N | Not A | O | B | | | B |
| 29 | 1,41 | 2,83 | 0,59 | B | Not A | B | B | | | B |
| 30 | 2,29 | 2,31 | 0,58 | N | Not A | B | B | | | B |

| A | B | N |
|------|-------|-------|
| 1 | 22 | 7 |
| 3,3% | 73,3% | 23,3% |

| A | B | N |
|------|-------|------|
| 1 | 27 | 2 |
| 3,3% | 90,0% | 6,7% |

| A | B | N |
|------|-------|------|
| 2 | 28 | 0 |
| 6,7% | 93,3% | 0,0% |