

348-93

34



ОБЪЕДИНЕННЫЙ
ИНСТИТУТ
ЯДЕРНЫХ
ИССЛЕДОВАНИЙ
ДУБНА

P10-93-348

В.В.Иванов

МНОГОМЕРНЫЙ АНАЛИЗ ДАННЫХ
НА ОСНОВЕ ИНТЕГРАЛЬНЫХ
НЕПАРАМЕТРИЧЕСКИХ СТАТИСТИК
И МНОГОСЛОЙНЫХ ПЕРЦЕПТРОНОВ

Направлено в «Nuclear Instruments and Methods»

1993

Введение

Основная цель обработки экспериментальных данных - извлечение интересующих исследователя событий (сигнальных событий) из совокупности данных, получаемых в эксперименте. Чаще всего для их отделения от фоновых событий используются обрезания по некоторой чувствительной к исследуемым эффектам переменной x . Простейший классификатор можно представить в виде пороговой функции $\Theta(x - x')$, которая равна 0 при $x < x'$ (область фоновых событий) и 1 при $x > x'$ (область сигнальных событий).

Барлоу несколько усовершенствовал этот алгоритм [1], предложив присваивать рассматриваемому событию вес, вычисляемый из соотношения

$$w(x) = \frac{s(x)}{s(x) + A \cdot b(x)} \quad \text{с } A = \frac{N_b}{N_s},$$

где $s(x)$ и $b(x)$ - ожидаемые распределения сигнальных и фоновых событий, а N_b и N_s - количество событий соответствующего сорта.

В том случае, когда событие характеризуется более чем одной переменной процедура построения многомерного классификатора не столь тривиальна. Интересное решение этой проблемы для произвольного числа независимых переменных было предложено Фишером [2] еще в 1936 году. Однако для идентификации событий, принадлежащих, в частности, двум разным распределениям необходимо предварительно оценить их параметры, что не всегда доступно в условиях реального эксперимента. Позднее для анализа многомерных распределений были развиты такие эффективные методы, как кластерный анализ [3], метод принципиальных компонент [4] и др.

В работах [5, 6] исследованы новые интегральные непараметрические статистики и построены одно- и двусторонние критерия согласия. На основе последних разработаны статистические методы выделения маловероятных событий на фоне от доминирующих процессов для случаев, когда анализируемые распределения представляют собой одновременные измерения одних и тех же случайных величин несколькими детекторами экспериментальной установки [7, 8]. Для применения этих методов необходимо знание параметров только доминирующего распределения.

В последнее время широкое применение в методике физического эксперимента находят искусственные нейронные сети, которые весьма успешно используются для классификации многомерных событий [9, 10]. Исследования показывают, в

частности, что такие классификаторы приближаются по своей эффективности к статистическим дискриминаторам [11].

В настоящей работе дается краткая информация об интегральных непараметрических критериях согласия и нейронносетевых классификаторах, проводится сравнительный анализ их мощностей и даются рекомендации по их совместному использованию.

1 Интегральные непараметрические критерии согласия

В задачах анализа данных физики частиц высоких и промежуточных энергий успешно применяются непараметрические критерии согласия на основе статистики ω_n^2 [12, 7], а также не столь известной статистики ω_n^3 [13].

Распределения ω_n^2 и ω_n^3 не зависят от F , поскольку их статистики могут быть записаны в виде интеграла от k -ой степени эмпирического процесса $v_n(t)$ [6]:

$$\omega_n^k = n^{\frac{k}{2}} \int_{-\infty}^{\infty} [S_n(x) - F(x)]^k dF(x) = \int_0^1 v_n^k(t) dt, \quad k = 1, 2, 3, \dots,$$

где $F(x)$ - теоретическая, а $S_n(x)$ - эмпирическая функции распределения случайной величины x , n - объем выборки. Это свойство позволяет изучать разнообразные распределения исследуемых величин, используя при этом общие свойства соответствующих критериев согласия.

Для практических целей удобно использовать обобщенное алгебраическое представление статистик ω_n^k , приведенное в работе [6]

$$\omega_n^k = -\frac{n^{\frac{k}{2}}}{k+1} \sum_{i=1}^n \left\{ \left[\frac{i-1}{n} - F(x_i) \right]^{k+1} - \left[\frac{i}{n} - F(x_i) \right]^{k+1} \right\} \quad (1)$$

В работе [6] даны также таблицы процентных точек для $k = 1, 2, \dots, 5$ при малых объемах эмпирической выборки $n = 1, 2, \dots, 10$.

Обычно критерии согласия используются для проверки гипотез о виде неизвестной функции распределения на основании анализа выборок из генеральной совокупности исследуемой случайной величины [14]. Критерии ω_n^k удобны тем, что с их помощью можно анализировать непосредственно наблюдаемые в каждом испытании значения рассматриваемой случайной величины, что позволяет исследовать каждую выборку (событие) на соответствие заранее известному распределению.

В работах [7, 8] на основе критериев ω_n^k была разработана следующая процедура выделения сигнальных событий:

а) анализируемые спектры преобразуются таким образом, чтобы вклады доминирующих распределений (чаще всего это распределения от фоновых событий) можно было аппроксимировать одной и той же функцией распределения $F_i(x)$;

б) каждая выборка, составленная из величин, одновременно принадлежащих преобразованным спектрам, исследуется на соответствие гипотезе, заданной функцией $F_i(x)$ при помощи критерия согласия ω_n^k ; при этом сигнальным событиям, не согласующимся с нулевой гипотезой, отвечают большие (по абсолютной величине) значения статистики ω_n^k (см. выражение (1)), что позволяет выделить их в критической области;

в) выделенные в критическую область события можно подвергнуть повторному анализу согласно п.п. а) и б) с той лишь разницей, что теперь в допустимую область отбираются (используя соответствующую функцию распределения $F_i(x)$) сигнальные события, а попавшие в критическую область - отбрасываются; тем самым достигается дополнительное исключение фоновых событий из исследуемых спектров.

В соответствии с указанной процедурой обработки был выполнен анализ данных ряда экспериментов.

В работе [7] был развит метод выделения маловероятных событий на основе критерия согласия ω_n^2 (критерий Смирнова-Крамера-Мизеса), который использовался для случая, когда анализируемые распределения представляют собой одновременные измерения времени пролета двумя различными системами. Анализ спектров масс вторичных частиц от фрагментации высокоэнергетических дейтронов на ядрах мишени позволил успешно выделить редкие события, связанные с образованием вторичных дейтронов.

В работе [8] на основе критерия согласия ω_n^3 предложен метод идентификации релятивистских заряженных частиц по измерениям ионизационных потерь одновременно в нескольких детекторах экспериментальной установки. Эта методика использовалась для идентификации вторичных заряженных частиц от фрагментации α -частиц высокой энергии, что позволило надёжно разделить события, в том числе и редкие, связанные с образованием различных одно- и дувзарядных частиц.

Для готовящихся на ускорителе *COSY (Julich, Germany)* опытов по изучению процессов подпорогового рождения K^+ путём численного эксперимента показано, что надёжное выделение редких K^+ -мезонов из спектров масс вторичных частиц может быть выполнено с использованием статистической процедуры, основанной на критерии согласия ω_n^3 [15].

2 Искусственные нейронные сети

Искусственные нейронные сети представляют собой новую модель параллельных распределённых вычислений, основанную на эмуляции характерных признаков живых нейронных сетей [16]. Их основу составляет элементарный вычислительный узел - нейрон, предназначенный для обработки подаваемой на входы информации и выработки выходного сигнала. Нейронная сеть состоит из большого числа связанных между собой нейронов. Связям между отдельными нейронами отвечают определенные веса, которые вычисляются в процессе настройки на решаемую задачу. По топологии связей между нейронами сети делятся на однослойные [17, 18] и многослойные. Одновременное участие в обработке большого числа нейронов позволяет максимально распараллелить вычислительный процесс и, соответственно, обеспечить высокую скорость анализа данных.

В последнее время искусственные нейронные сети находят широкое применение в методике физического эксперимента [9, 10]. Они используются как для построения интеллектуальных триггеров высокого уровня, позволяющих проводить эффективную дискриминацию фоновых событий [19, 20] в реальном времени эксперимента, так и в задачах обработки экспериментальных данных, таких как распознавание траекторий частиц [21, 22], восстановление масс, идентификация первичных и/или вторичных вершин [23, 24, 25] и т.д.

Наибольшее распространение получили многослойные feed forward сети - мно-

гослойный перцептрон [9, 11]. Такая сеть состоит из нескольких слоев: входного, на который подаются исследуемые данные, одного или нескольких скрытых и выходного, с которого снимается результат.

Архитектура сети представлена на рис.1.

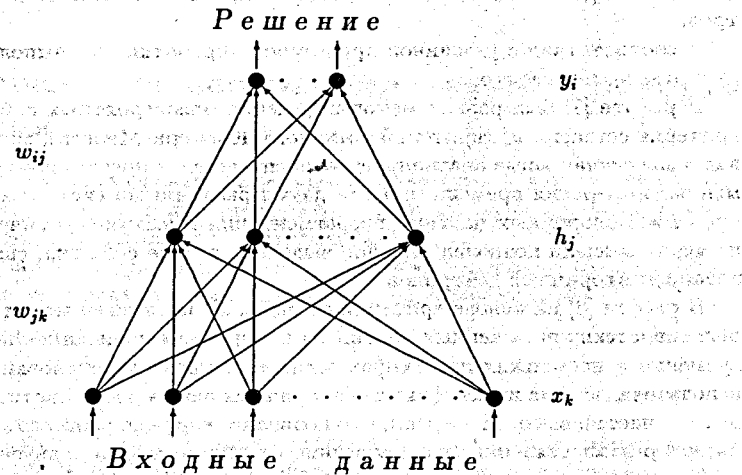


Рис.1. Архитектура многослойного перцептрона с одним скрытым слоем

Здесь через x_k , h_j и y_i обозначены, соответственно, входные, скрытые и выходные нейроны; w_{jk} – веса связей входных нейронов со слоем скрытых, а w_{ij} – веса связей скрытых нейронов с выходными нейронами. На входы скрытых и выходных нейронов подаются соответственно сигналы $a_j = \sum_k w_{jk} x_k$ и $a_i = \sum_j w_{ij} h_j$. Величины выходных сигналов с этих нейронов определяются из соотношений

$$h_j = g(a_j/T) + \theta_j,$$

$$y_i = g(a_i/T) + \theta_i,$$

где $g(a, T)$ – переходная функция, T – "температура", задающая величину ее наклона, а θ – порог соответствующего узла. В качестве $g(a, T)$ чаще всего используются функции сигмоидного типа, например, вида

$$g(a, T) = \frac{1}{2} [1 + \tanh(\frac{a}{T})]. \quad (2)$$

Настройка сети на конкретную задачу реализуется путём её обучения, для чего чаще всего используется алгоритм back-propagation [26]. В процессе обучения слой скрытых нейронов выполняет задачу настройки на изучаемые распределения, что реализуется коррекцией весов w_{jk} ; таким образом, чтобы для подаваемых на вход сети выборок переменных $\vec{x}^{(p)}$ ($p = 1, \dots, M$ – число тренировочных итераций) выходные сигналы $\vec{y}^{(p)}$ совпадали с целевым значением $\vec{t}_i^{(p)}$. Эта задача решается путём минимизации функционала

$$E = \frac{1}{2} \sum_p \sum_i (y_i^{(p)} - t_i^{(p)})^2,$$

получившего название функции ошибок (или функции энергии сети).

Известно (см. [9]), что на основе такой сети можно построить многомерный классификатор, однако его мощность существенно зависит от выбора входных данных. Примеры использования многослойного перцептрона в задачах анализа многомерных распределений в экспериментальной физике высоких энергий хорошо представлены в обзорах (см. [9, 10]).

3 Анализ данных с помощью критериев ω_n^k и многослойных перцептронов

Сравнение мощностей указанных классификаторов проводилось для двух задач, в которых многомерные события генерировались методом Монте-Карло.

3.1 Первый пример

В первом случае моделировались потери энергии на ионизацию частицами разных сортов одновременно в нескольких детекторах экспериментальной установки. В качестве примера рассматривалась приведённая в [27] задача по разделению протонов и заряженных пионов с энергией 100 ГэВ в потоках космических лучей.

Моделирование событий, каждое из которых представляло собой набор случайных величин – ионизационных потерь энергии, испытываемых пионами или протонами в нескольких детекторах установки, проводилось следующим образом. Вначале разыгрывался сорт регистрируемой частицы, при этом отношение вкладов пионов и протонов полагалось равным $\pi^+ : p = 1 : 4$ [27]. Затем генерировалось n случайных величин ΔE_i , $i = 1, 2, \dots, n$ (n – число детекторов в установке равнялось 6) в соответствии с законом распределения ионизационных потерь для частиц рассматриваемого сорта. Общее число событий было взято равным 10000. На рис.2 сплошной кривой представлен суммарный результат такого моделирования для отдельного детектора, вклады от протонов и пионов выделены штриховкой.

При использовании критерия ω_n^k в качестве нулевой гипотезы выбиралась функция распределения Ландау $\phi(\lambda)$ (см. [8]). Элементами эмпирической выборки служили величины λ :

$$\lambda_i = \frac{\Delta E_i - \Delta E_{mp}^i}{\xi_i}, \quad i = 1, 2, \dots, n, \quad (3)$$

где ΔE_i – величина потери энергии в i -м счетчике, ΔE_{mp}^i – приведённое значение наиболее вероятной потери энергии (см. разд.2 в [8]), $\xi_i \approx \frac{1}{4} \text{FWHM}$ для распределения ионизационных потерь протона в отдельном детекторе.

Так как известно (см. [8]), что использование односторонних критериев для рассматриваемой задачи более предпочтительно, а также то, что с ростом k растёт мощность соответствующего критерия [6], то для каждого события вычислялись величины ω_n^k по формуле (1) при $n = 6$ и максимальной степени $k = 5$, для которой имеются таблицы процентных точек [6].

На рис.3 приведено распределение случайных величин ω_6^5 , полученное в результате обработки сгенерированных событий. "Чистая" гистограмма образована протонными событиями; распределение, относящееся к пионам, выделено штриховкой и располагается в основном в области больших ω_6^5 . Одновременно подсчитывались вероятности регистрации протонных и пионных событий в зависимости от

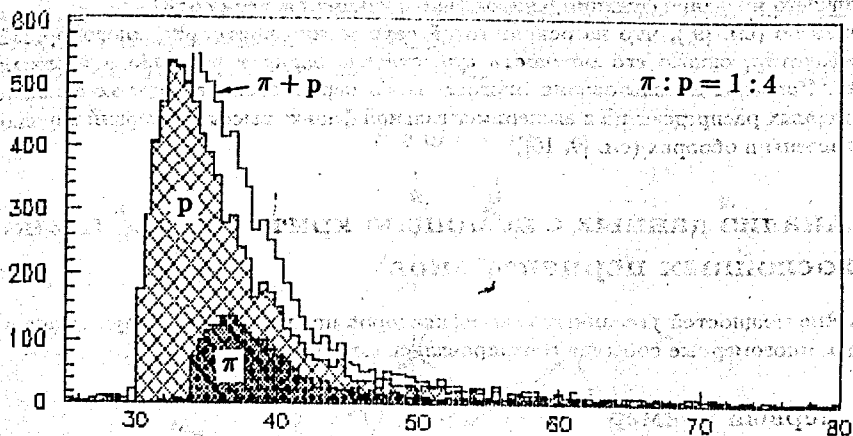


Рис.2. Суммарное распределение потерь энергии для 100 ГэВ-иных протонов и пионов в отдельном детекторе (в кэВ); вклады от протонов и пионов выделены штриховкой.

величины y_0 граничного значения статистики ω_0^5 . На рис.4 представлены кривые зависимости величины накопленной вероятности $F(y) = \Pr\{y < y_0\}$ для пионных и зависимости $1 - F(y)$ для протонных событий.

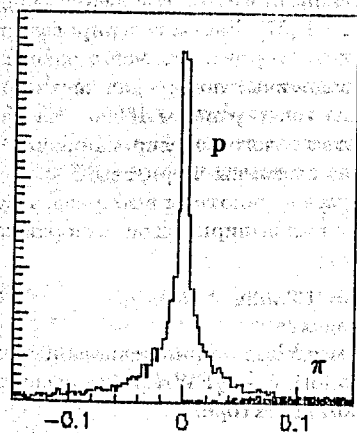


Рис.3. Распределение случайных величин ω_0^5 , полученное в результате обработки сгенерированных событий; "чистая" гистограмма отвечает протонам, распределение от пионов выделено штриховкой.

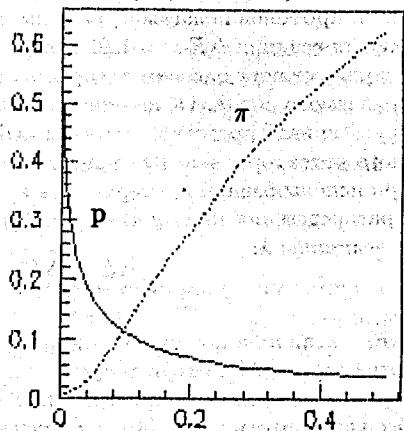


Рис.4. Кривые зависимости величины накопленной вероятности $F(y) = \Pr\{y < y_0\}$ для пионных и зависимости $1 - F(y)$ для протонных событий; критерий ω_0^5 .

имели одну и ту же передаточную функцию вида (2). Полагалось, что для протонных событий сигнал на выходе сети должен быть равен -1 , а для пионных $+1$. Попытка использовать в качестве входных данных выборку величин потерь энергии ΔE_i , $i = 1, \dots, n$ не дала положительного результата. Поэтому, также как и для критерия ω_n^5 , в качестве входной информации на сеть подавалась упорядоченная выборка величин λ_i , $i = 1, \dots, n$, вычисленных согласно соотношению (3).

Все сгенерированные события – около 2000 пионных и 8000 протонных событий с соответствующими признаками – были перемешаны и разбиты на две одинаковые по количеству событий группы. Первая группа использовалась для обучения сети, а вторая – для оценки её способности классифицировать события. Идентификация событий осуществлялась по величине выходного сигнала: если она не превышала некоторое заданное пороговое значение, то рассматриваемое событие считалось протонным, в противном случае – пионным.

В среднем выполнялось до 20 тренировочных эпох, хотя примерно после 10 эпохи сеть приходила в стабильное состояние. В каждой последующей эпохе на вход сети подавались одни и те же данные, а по её завершении проводилась коррекция весов межнейронных связей. По окончании обучения сети фиксировались веса и пороги-смещения нейронов.

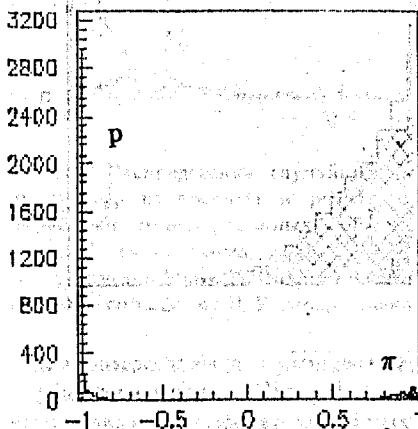


Рис.5. Спектры величины выходных сигналов сети, полученные в результате обработки сгенерированных событий: "чистая" гистограмма – протонные события, штрихованная гистограмма – пионные события.

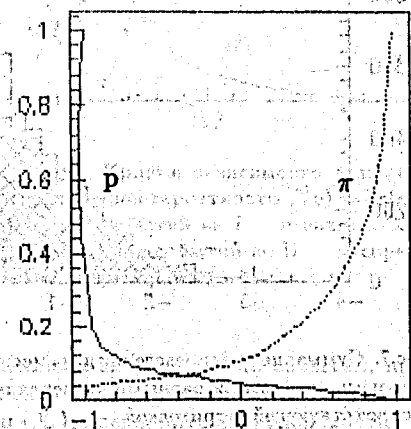


Рис.6. Кривые зависимости величины накопленной вероятности $F(y) = \Pr\{y < y_0\}$ для пионных и зависимости $1 - F(y)$ для протонных событий; нейронная сеть.

При тестировании сети на её вход подавалась смесь событий от обеих частиц. На рис. 5 приведены спектры величин выходных сигналов сети, полученные в результате обработки сгенерированных событий: "чистая" гистограмма – протонные события, штрихованная гистограмма – пионные события. При этом подсчитывалась вероятность распознавания событий разного сорта из смеси, которая для случая, когда в качестве граничной точки был выбран 0, составила

93,5%. Кроме того, также как для критерия ω_3^2 , подсчитывались вероятности распознавания протонных и пионных событий в зависимости от величины порога y_p . На рис.6 представлены кривые зависимости величины накопленной вероятности $F(y) = \Pr\{y < y_p\}$ для пионных и зависимости $1 - F(y)$ для протонных событий.

3.2 Второй пример

Во втором примере решалась задача распознавания событий, представляющих собой выборки $x_i, i = 1, \dots, n$ объёмом $n = 6$ из двух одномерных гауссианов с одинаковыми средними значениями ($a_1 = a_2 = 0$) и разными среднеквадратичными отклонениями ($\sigma_1 = 1, \sigma_2 = 0,3$). Как и в первом примере, число разыгрываемых событий из первого распределения соотносилось с числом событий из второго распределения как 4:1. Общее число сгенерированных событий равнялось 10000. На рис.7 "чистой" гистограммой представлен суммарный результат такого моделирования; вклады от I и II гауссианов выделены соответствующей штриховкой.

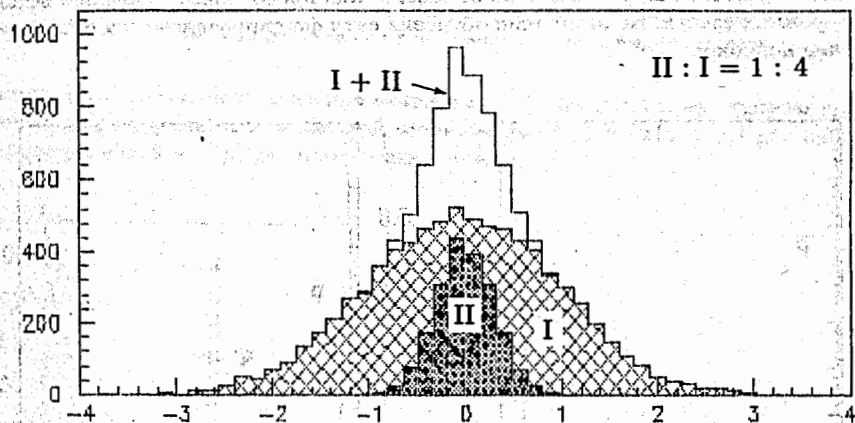


Рис.7. Суммарное распределение для смеси выборок из двух одномерных гауссианов с одинаковыми средними и разными дисперсиями; вклады от I и II гауссианов выделены соответствующей штриховкой

Для этой задачи более предпочтительно использование двусторонних критериев ω_k^2 с чётными степенями $k: k = 2, 4, \dots$. Кроме того, так как с ростом k растёт мощность соответствующего критерия [6], то была взята статистика с максимальной степенью $k = 4$, для которой имеются таблицы процентных точек [6]. В качестве нулевой гипотезы в формуле (1) использовалось стандартное гауссовское распределение $N(0, 1)$ (см. [7]). Элементы эмпирической выборки получались из набора случайных переменных $\{x_i\}$ путём замены переменных:

$$X_i = \frac{x_i - a_1}{\sigma_1}, \quad i = 1, \dots, n. \quad (4)$$

Такое преобразование приводит к тому, что выборки $\{X_i\}$ из первого распределения описываются одной и той же функцией распределения, которая соответствует

плотности $N(0, 1)$, значения же функции распределения для конкурирующего распределения имеют тенденцию группироваться в области I.

На рис.8 приведены распределения случайных величин ω_6^4 , полученные в результате обработки сгенерированных событий; "чистая" гистограмма отвечает событиям от I распределения, штрихованная гистограмма – событиям от II распределения. На рис.9 представлены кривые зависимости величины накопленной вероятности $F(y) = \Pr\{y < y_p\}$ для событий из I-го и зависимости $1 - F(y)$ для событий из II-го распределений.

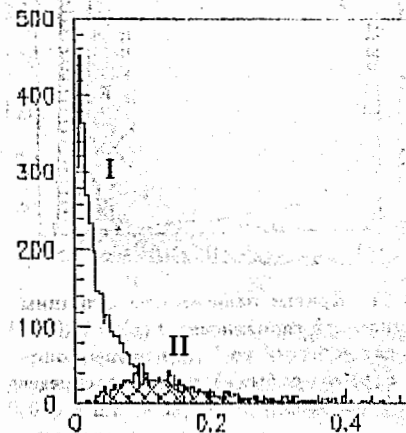


Рис.8. Распределения случайных величин ω_6^4 , полученные в результате обработки сгенерированных событий; "чистая" гистограмма – события от I распределения; штрихованная гистограмма – события от II распределения

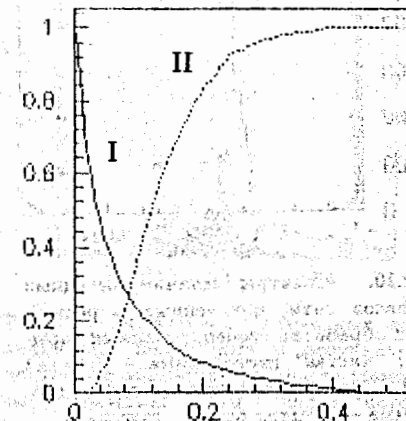


Рис.9. Кривые зависимости величины накопленной вероятности $F(y) = \Pr\{y < y_p\}$ для событий из I и зависимости $1 - F(y)$ для событий из II распределений; критерий ω_6^4

По аналогии с первой задачей была выполнена классификация событий, принадлежащих различным гауссианам, с помощью нейронной сети. При этом на её вход подавались выборки случайных величин $\{X_i\}$, предварительно упорядоченных в порядке возрастания. На рис.10 приведены спектры величин выходных сигналов сети, полученные в результате обработки сгенерированных событий: "чистая" гистограмма – события от I гауссиана, штрихованная гистограмма – события от II гауссиана. На рис.11 представлены соответствующие кривые зависимости величины накопленной вероятности $F(y) = \Pr\{y < y_p\}$ для событий I сорта и зависимости $1 - F(y)$ для событий II сорта.

4 Анализ результатов

4.1 Анализ результатов первого примера

Из графиков, приведённых на рис.4 и 6, видно, что нейронная сеть лучше классифицирует протонные и пионные события. Детальный анализ этих графиков показал,

ошибки первого и второго рода в месте пересечения кривых составляют для критерия ω_0^5 11,4% и 11,5%, а для нейронной сети 6,8% и 6,9%.

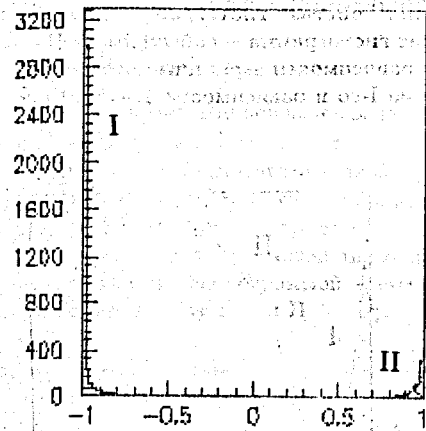


Рис.10. Спектры величин выходных сигналов сети, полученные в результате обработки сгенерированных событий: "чистая" гистограмма - события от I гауссиана, штрихованная гистограмма - события от II гауссиана

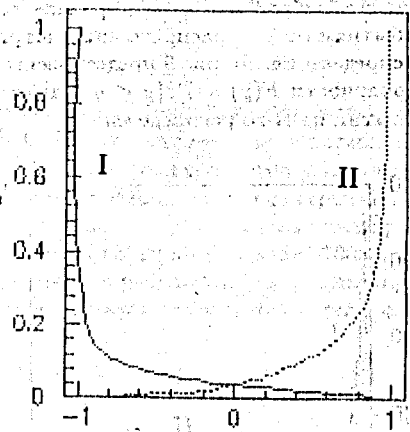


Рис.11. Кривые зависимости величины накопленной вероятности $F(y) = \Pr\{y < y_j\}$ для событий из I и зависимости $1 - F(y)$ для событий из II распределений: нейронная сеть

Однако, как отмечалось в пункте в) в приведённой выше процедуре обработки, после того как выбрана критическая граница, можно провести повторную обработку событий выделенных в критической области, используя в качестве нулевой гипотезы распределение ионизационных потерь для пионов. Очевидно, что критическую границу целесообразно выбрать таким образом, чтобы потеря пионных событий была минимальной, а вклад протонных событий в критическую область не был слишком заметным.

Критическая граница была взята равной 0,045. Общее число событий, попавших в критическую область, составило 3099; из них 1951 пионных и 1148 протонных событий. Анализ распределений потерь энергии для пионов и протонов, для которых случайная величина $\omega_0^5 \geq 0,045$, показал, что часть протонных событий можно выделить из смеси, если взять в качестве порога величину потери энергии $\Delta E = 33,5$ МэВ. При этом "теряется" 63 пионных события, однако из рассматриваемой выборки "выбрасывается" 619 протонных событий. В результате ошибки первого и второго рода составили 6,6% и 6,4%, что близко к тому, что получено с помощью нейронной сети, и практически совпадает с результатом, полученным в работе [8] с помощью метода отношения функций правдоподобия - наиболее мощного метода в случае простых гипотез.

На рис.12 представлено суммарное распределение потерь энергии протонов и пионов в отдельном детекторе для событий с $\omega_0^5 \geq 0,045$ и $\Delta E > 33,5$ МэВ; вклады от протонов и пионов выделены штриховкой. Результат повторной обработки этих событий, при которой в качестве нулевой гипотезы взято распределение ионизационных потерь для пионов, показан на рис.13. Из него видно, что дальнейшее улучшение разделения протонов и пионов практически невозможно.

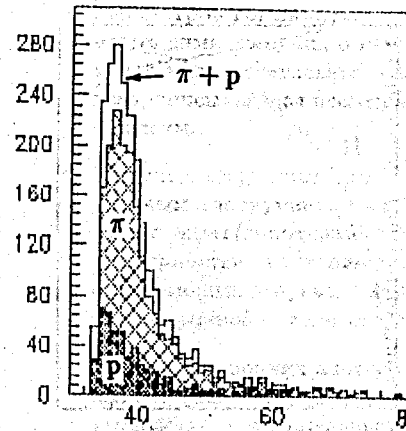


Рис.12. Суммарное распределение потерь энергии протонов и пионов в отдельном детекторе для событий с $\omega_0^5 \geq 0,045$ и $\Delta E > 33,5$ МэВ; вклады от протонов и пионов выделены штриховкой.

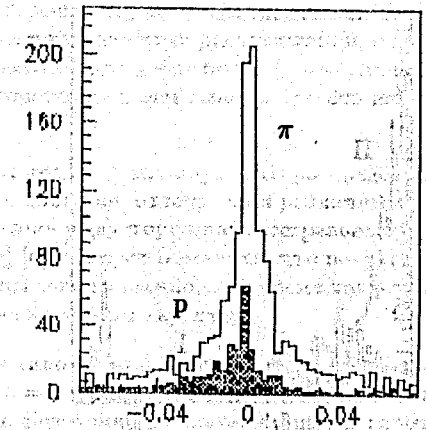


Рис.13. Гистограммы случайных величин ω_0^5 , полученные в результате повторной обработки рассматриваемой смеси событий: пионы - чистая гистограмма, протоны - штрихованная гистограмма

4.2 Анализ результатов второго примера

Из графиков, приведённых на рис.9 и 11, видно, что нейронная сеть существенно лучше идентифицирует события, отвечающие I или II гауссиану. Анализ показал, что ошибки первого и второго рода в месте пересечения кривых составляют для критерия ω_0^4 26,7% и 25,3%, а для нейронной сети 4,0% и 4,1%.

Выполним повторную обработку событий выделенных в критической области, используя в качестве нулевой гипотезы распределение, отвечающее второму гауссиану. Если выбрать границу равной 0,035, то общее число событий, попавших в критическую область, составит 3009; из них 997 событий от II-го и 2012 от I гауссианов соответственно¹. На рис.14 приведены распределения случайных величин ω_0^4 для событий от II гауссиана полностью ("чистая" гистограмма) и для части событий от I гауссиана (на рисунке она штрихована), попавших в указанную область. Кроме того, на рис.15 представлены соответствующие кривые зависимости величины накопленной вероятности $F(y) = \Pr\{y < y_j\}$ для событий II сорта и зависимости $1 - F(y)$ для событий I сорта.

Если теперь выбрать в качестве критической границы значение равное 0,2 и подсчитать число неидентифицированных (с помощью критерия ω_0^4) событий, то окажется, что доля таких событий от I гауссиана составляет 8,4%, а от II

¹Общее число событий в контрольной выборке в этом случае вдвое было меньше нежели в первом примере.

- 13,0%. Таким образом, хотя процедура повторного применения критерия ω_n^k заметно улучшила результаты, полученные на первом этапе, однако в этом случае не удалось приблизиться к результату, достигнутому с помощью нейронной сети.

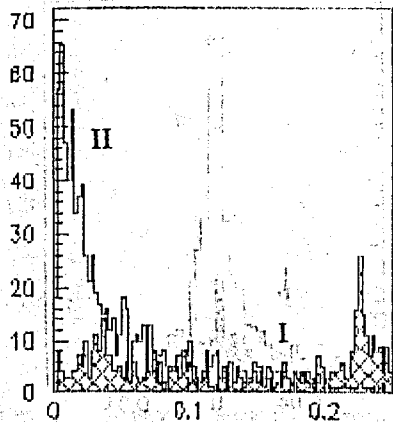


Рис.14. Распределения случайных величин ω_n^k для событий от II гауссиана полностью ("чистая" гистограмма) и для части событий от I гауссиана (на рисунке она заштрихована), попавших в указанную область

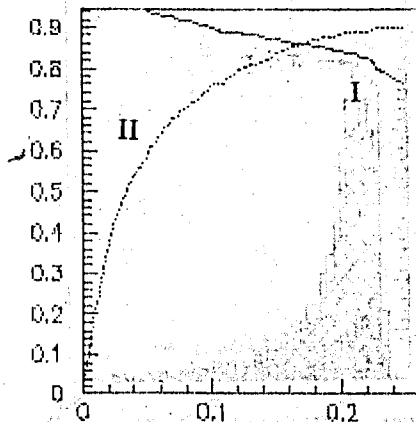


Рис.15. Кривые зависимости величины накопленной вероятности $F(y) = \Pr\{y < y_0\}$ для событий из II и зависимости $1 - F(y)$ для событий из I распределений: повторное применение критерия ω_n^k

Заключение

Проведён сравнительный анализ мощностей многомерных классификаторов на основе интегральных непараметрических критериев согласия ω_n^k и многослойных перцептронов для случаев, когда анализируемые распределения представляют собой одновременные измерения одних и тех же физических величин несколькими детекторами экспериментальной установки².

Результаты настоящей работы можно просуммировать следующим образом:

1. Многослойный перцептрон слабо классифицирует события, если в качестве входных данных используются многомерные выборки случайных величин из исходных распределений. Наличие дополнительных измерений не приводит практически к увеличению вероятности распознавания сетью событий, отвечающих различным гипотезам.
2. Эта же сеть обеспечивает мощность, близкую к предельной, если классификация событий проводится в пространстве эффективных переменных. В качестве таких переменных целесообразно использовать выборки случайных

²Рассмотренные примеры являются типичными для широкого класса задач экспериментальной физики частиц промежуточных и высоких энергий (см., например, работы [7, 8, 15, 27, 28]).

величин из анализируемых распределений, нормированных согласно соотношений (3) или (4), и упорядоченных в порядке возрастания; такое преобразование исходных переменных дало прекрасные результаты в обоих примерах.

3. Для перехода от исходной выборки к новым переменным необходимо знание параметров доминирующего распределения. Кроме того, на этапе обучения необходимы сведения о параметрах альтернативных распределений, образующих экспериментальный спектр. Очевидно, что чем более реалистична модель, описывающая исследуемые распределения, тем выше мощность нейронной сети.
 4. В случае, когда известны параметры только доминирующего распределения, удобным инструментом для селекции событий, отвечающих различным распределениям (гипотезам), являются одно- и двусторонние интегральные непараметрические критерии согласия ω_n^k [6]. Следует отметить, что получаемые с их помощью результаты обоснованы количественно, в то время как результат нейронной сети носит только качественный характер.
 5. Для односторонних альтернативных гипотез критерии согласия ω_n^k с нечётными k (см., примеры с $k = 3$ [7] и $k = 5$ - настоящая работа) дают результаты близкие к предельным; в случае двусторонних альтернативных гипотез мощность критериев согласия ω_n^k с чётными k несколько уступает мощности, достигаемой нейронной сетью.
 6. Критерии ω_n^k удобны тем, что путём их последовательного применения можно выделять вклады любого числа парциальных распределений из наблюдаемого в эксперименте результирующего спектра. Это позволяет, в частности, оценив параметры распределений, образующих смесь, в случае необходимости, воспользоваться дополнительно нейронной сетью.
 7. Так как имеются аппаратные реализации нейронных сетей рассматриваемого типа [29], обладающих высоким быстродействием, а процедура перехода к новым переменным может быть выполнена на быстрых спецпроцессорах, то на этой основе можно строить высокопроизводительные классификаторы многомерных событий, обладающие мощностью близкой к предельной.
- Таким образом, совместное использование интегральных непараметрических статистик и многослойных перцептронов даёт исследователям эффективный инструмент для анализа многомерной информации, как в реальном времени эксперимента, так и в процессе обработки накопленной информации.

Литература

- [1] R.Barlow: J.Comput.Physics, 72, 1987, p.202.
- [2] R.A.Fisher: Ann. Eugen.7 (1936) p.179.
- [3] Th.Naumann and H.Schiller: "Multidimensional Data Analysis", in: "Formulae and Methods in Experimental Data Evaluation", vol.3 (European Physical Society, Geneve, 1983).
- [4] R.Brun et al.: LINTRA, CERN Program Library E 5002.
- [5] P.V.Zrelov, V.V.Ivanov: "Test Statistics $\omega_n^3 = n^{\frac{1}{2}} \int_{-\infty}^{\infty} [S_n(x) - P(x)]^3 dP(x)$ and its Main Properties". Transactions of the Eleventh Prague Conference on Information Theory, Statistical Decision Functions, Random Processes. Prague, from August 27 to 31, 1990. ACADEMIA Publishing House of the Czechoslovak Academy of Sciences, Prague 1992, p.487 - 498.
- [6] П.В.Зрелов, В.В.Иванов: "Интегральные непараметрические статистики $\omega_n^k = n^{\frac{k}{2}} \int_{-\infty}^{\infty} [S_n(x) - P(x)]^k dP(x)$ и их основные свойства. Алгебраический вид, функции распределения и критерии согласия". Сообщения ОИЯИ, P10-92-461, Дубна, 1992.
- [7] П.В.Зрелов, В.В.Иванов: "Метод выделения маловероятных событий с помощью критерия согласия Смирнова-Крамера-Мизеса". Препринт ОИЯИ, P10-86-812, Дубна, 1986.
- [8] P.V.Zrelov and V.V.Ivanov: "The Relativistic Charged Particles Identification Method Based on the Goodness-of-Fit ω_n^3 -Criterion". Nucl. Instr. and Meth. in Phys. Res., A310 (1991) 623-630.
- [9] B.Denby: "Tutorial on Neural Networks Applications in High Energy Physics: 1982 Perspective". In Proc. of the Second International Workshop on "Software Engineering, Artificial Intelligence and Expert Systems in High Energy Physics". January 13-18, 1992 L'Agelaide France-Telecom La Londe-les-Maures (France). New Computing Techniques in Physics Research II, edited by D.Perret-Gallix, World Scientific, 1992, p.287.
- [10] C.Peterson: "Neural Networks in High Energy Physics - Algorithms and Results". Ibidem, p.327.
- [11] F.Fogelman Soulie: "Neural Networks for Patterns Recognition: Introduction and Comparison to Other Techniques". Ibidem, p.277.
- [12] T.Ludlam et al.: Phys.Rev.D, 1973, v.8, No.5, p.408; Phys.Rev.D, 1977, v.16, No.1, p.100.
- [13] P.V.Zrelov and V.V.Ivanov: "Experimental Analysing Methods Based on the Goodness-of-Fit Criterion ω_n^3 ". In: "Computing in High Energy Physics'91, 1991, p.503-505.
- [14] "Статистические методы в экспериментальной физике". Пер. с англ. Под ред. А.А.Тяпкина. Атомиздат, М.: 1976.
- [15] П.В.Зрелов, В.В.Иванов, В.И.Комаров, А.И.Пузынин, А.С.Хрыкин: "Моделирование эксперимента по изучению процессов подпорогового рождения K^+ -мезонов". Препринт ОИЯИ, P10-92-369, Дубна, 1992; "Математическое моделирование", т.4, №11, 1993, с.56-74.
- [16] R.Lippman: "An Introduction to Computing with Neural Nets", IEEE ASSP Magazine 4 (1987).
- [17] J.J.Hopfield and D.W.Tank: Science, 233 (1986) 625.
- [18] А.Ю.Лоскутов и А.С.Михайлов: "Введение в синергетику", М.: "Наука", (1990), с. 233.
- [19] L.Gupta, A.Upadhye, B.Denby and S.R.Amendolia: "Neural Network Trigger Algorithms for Heavy Quarks Event Selection in Fixed Target High Energy Physics Experiments". Fermilab-Pub 91/117, Submitted to Pattern Recognition.
- [20] P.Ribarics et al.: "Neural Network Trigger in the H1 Experiment". In Proc. of the Workshop "Neural Networks: From Biology to the High Energy Physics", Ella International Physics Center, Isola d'Ella, Italy, June 5-14, 1991, ETS Editrice, Pisa.
- [21] M.Ohlsson, C.Peterson and A.Yuille: "Track Finding with Deformable Templates - The Elastic Arms Approach". Lund University Preprint LU TP 91-27, November 1991, Lund, Sweden. Submitted to Computer Physics Communications.
- [22] M.Gyulassy and H.Harlander: "Elastic Tracking and Neural Networks Algorithms for Complex Pattern Recognition". Computer Physics Communications, 66, 1991, p.31.
- [23] L.Lonnblad, C.Peterson and T.Rogvaldsson: "Mass Reconstruction with Neural Networks". Lund University Preprint LU TP 91-25, October 1991, Lund, Sweden. Submitted to Physics Letters B.
- [24] C.S.Lindsey: "Tracking and Vertex Finding in Drift Chambers with Neural Networks". In Proc. of the Workshop "Neural Networks: From Biology to the High Energy Physics", Ella International Physics Center, Isola d'Ella, Italy, June 5-14, 1991, ETS Editrice, Pisa.
- [25] C.S.Lindsey, B.Denby, H.Haggerty, K.Johns: "Real Time Finding in a Drift Chamber with a VLSI Neural Network". Fermilab-Pub 92/55, accepted for publication in Nucl. Instr. and Meth.a.
- [26] D.E.Rumelhart, G.E.Hinton, R.J.Williams: "Learning Internal Representations by Error Propagation" in D.E.Rumelhart & J.L.McClelland (Eds.), Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Vol. 1: Foundations. MIT Press (1986).
- [27] P.V.Ramaha Murty and G.D.Demeester: Nucl. Instr. and Meth., 1967, 56, p.93.

- [28] G.Basti et al.: "Automatic Redefinition of the Fuzzy Membership Function to Deal with High Fluctuating Phenomena in Neural Nets". INFN Roma II, Italy, April 12, 1993.
- [29] J.R.Hansen: "The Need for Neural Networks at LHC/SSC". In Proc. of the Second International Workshop on "Software Engineering, Artificial Intelligence and Expert Systems in High Energy Physics". January 13-18, 1992 L'Agelade France-Telecom La Londe-les-Maures (France). New Computing Techniques in Physics Research II, edited by D.Perret-Gallix, World Scientific, 1992, p.353.

Рукопись поступила в издательский отдел
17 сентября 1993 года.