

УДК 004.93'1:519.237:51-72

Об идентификации электронов и пионов с помощью многослойного перцептрона в детекторе переходного излучения эксперимента CBM

Т. П. Акишина, О. Ю. Дереновская, В. В. Иванов

*Лаборатория информационных технологий
Объединённый институт ядерных исследований
141980, Дубна, Московская область, Россия*

Рассмотрена задача идентификации электронов и пионов по их потерям энергии в детекторе переходного излучения (TRD) эксперимента CBM. Для идентификации частиц использовалась искусственная нейронная сеть — многослойный перцептрон, реализованный в пакетах JETNET и ROOT. В работе показано, что для получения корректных и сопоставимых результатов важно правильно выбрать структуру сети и приводятся соответствующие рекомендации. Для достижения приемлемого уровня подавления пионов требуется преобразование величин потерь энергии в слоях TRD к более «эффективным» переменным.

Ключевые слова: многомерные методы анализа данных, методы распознавания образов, эксперимент CBM, детектор переходного излучения TRD, искусственные нейронные сети.

1. Введение

Экспериментальная установка CBM (Compressed Baryonic Matter), создаваемая в ГСИ (Дармштадт, Германия) на ускорительном комплексе антипротонов и тяжёлых ионов FAIR (Facility for Antiproton and Ion Research), нацелена на изучение свойств сильно сжатой барионной материи, образующейся в ядро-ядерных соударениях при энергии пучка $8 \div 45$ ГэВ на нуклон [1, 2].

Набор детекторов и элементов установки CBM должен обеспечить: идентификацию электронов при условии подавления пионов на уровне 10^5 , идентификацию адронов с большим акцептансом, восстановление импульсов заряженных частиц с точностью $\sim 1\%$, восстановление первичных и вторичных вершин с точностью ~ 30 мкм, высокое пространственное разрешение координатных детекторов, малое мёртвое время, быстрый ответ и высокую скорость считывания информации с детекторов, высокую скорость триггера и системы сбора данных, устойчивость детекторов и электроники к радиационным условиям, устойчивость к дельта-электронам и т.д.

На рис. 1 представлена схема экспериментальной установки CBM. Внутри дипольного магнита расположена мишень и координатная трековая система (STS — Silicon Tracking System), содержащая 7 плоскостей из пиксельных и стриповых детекторов. STS совместно с дипольным магнитом используется для восстановления траекторий заряженных частиц и определения их импульсов. Черенковский детектор (RICH — Ring Imaging Cherenkov) и детектор переходного излучения (TRD — Transition Radiation Detector) должны обеспечить надёжную регистрацию электронов с импульсом выше 1 ГэВ/с. Детектор измерения времени пролёта (TOF — Time-of-flight), построенный на основе резистивных плоских камер (RPC — Resistive Plate Chambers), предназначен для идентификации адронов высокой энергии. Электромагнитный калориметр (ECAL — Electromagnetic Calorimeter) используется для идентификации электронов и фотонов.

Измерение чармония — одна из ключевых задач эксперимента CBM. Для того, чтобы зарегистрировать J/ψ мезон при его распаде по диэлектронному каналу,

Статья поступила в редакцию 22 октября 2009 г.

Авторы благодарят Лебедева С.А. за помощь в подготовке модельных данных и полезные обсуждения.

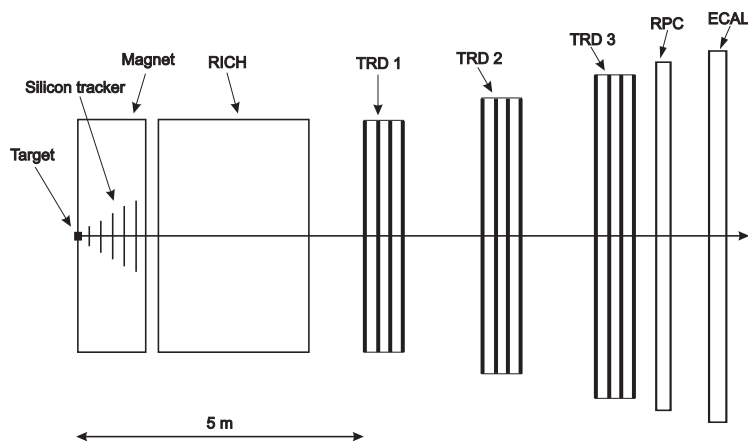


Рис. 1. Схема эксперимента CBM

необходимо решить главную задачу — надёжно идентифицировать электроны в условиях доминирующего фона от пионов. Одним из наиболее эффективных детекторов для решения указанной задачи является многослойный детектор переходного излучения TRD.

TRD должен обеспечить надёжную идентификацию электронов, высокий уровень подавления пионов, а также восстановление траекторий проходящих через него заряженных частиц. Требуемое подавление пионов должно быть не менее 100, пространственное разрешение около 200–300 мкм. Чтобы решить указанные задачи в условиях интенсивных потоков и высокой множественности частиц, необходима тщательная оптимизация детектора TRD.

Ранее в работе [3] нами была исследована возможность решения указанной задачи на основе искусственной нейронной сети, и было показано, что эта задача может быть решена с помощью многослойного перцептрона (МСП), входящего в пакет JETNET [4]. В эксперименте CBM для этих же целей используется МСП [5], реализованный в среде CBM ROOT [6, 7]. Сравнение эффективностей работы этих сетей дало заметно различающиеся результаты. В настоящей работе исследуются причины такого расхождения.

В работе [3] было также показано, что для обеспечения максимального уровня подавления пионов требуется преобразование величин потерь энергии в слоях TRD к более «эффективным» переменным. Здесь нами получена зависимость параметров данного преобразования от импульса регистрируемой частицы.

2. Основные принципы распознавания образов с помощью МСП

Многослойная прямоточная нейронная сеть — многослойный перцептрон — удобный инструмент для построения многомерных классификаторов, хотя скорость его обучения и мощность распознавания критически зависят от выбора входных данных.

МСП состоит из нескольких слоёв нейронов: слой входных нейронов, одного или нескольких слоёв скрытых нейронов и слоя выходных нейронов. Связи в таких сетях устанавливаются между нейронами соседних слоёв, а движение информации происходит в одном направлении: от входного слоя к выходному.

Проблема классификации состоит в определении принадлежности входных данных (представленных вектором) к одному или нескольким априори определённым классам. Анализируемые данные подаются на входной слой МСП, а сигнал, получаемый с выходного слоя нейронов, позволяет определить — к какому классу принадлежит анализируемый образец.

Основные принципы работы МСП и его применения к задаче классификации удобно рассмотреть на примере, взятом из статьи Б. Денби [8]. Пусть нам необходимо построить классификатор событий, который позволяет определить к какому классу, «а» или «b», относится анализируемый образец (рис. 2).

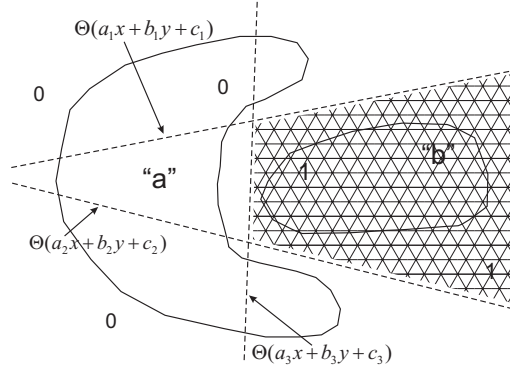


Рис. 2. Области распределения событий, отвечающих двум классам: «а» и «b»

Дискриминирующая функция, отвечающая решению этой задачи, имеет следующий вид:

$$D = \theta[a_1x + b_1y + c_1 + a_2x + b_2y + c_2 + a_3x + b_3y + c_3 - 2], \quad (1)$$

где пороговая функция $\theta(x - x')$ равна 0, если $x < x'$ и 1 для $x \geq x'$. Параметры a_i , b_i и c_i , $i = 1, \dots, 3$ выбраны таким образом, чтобы функция (1) принимала значение 1 в области, которая включает границы класса «b», и значение 0 — в остальной области.

На рис. 3 представлена схема дискриминатора, который реализован с помощью функции (1). На входы x и y подаются случайные переменные, которые соответствуют текущему событию. Их величины умножаются на коэффициенты a_i и b_i , $i = 1, \dots, 3$, которые отвечают весам связей, соединяющих входы с первым слоем пороговых дискриминаторов (ПД). На входы ПД также подаются пороги, которые отвечают за параллельный сдвиг, соответствующий разделяющей линии. Сигналы с выходов скрытых ПД умножаются на новые веса (которые в данном случае равны 1), и вместе с порогом, равным -2 , подаются на выходной ПД.

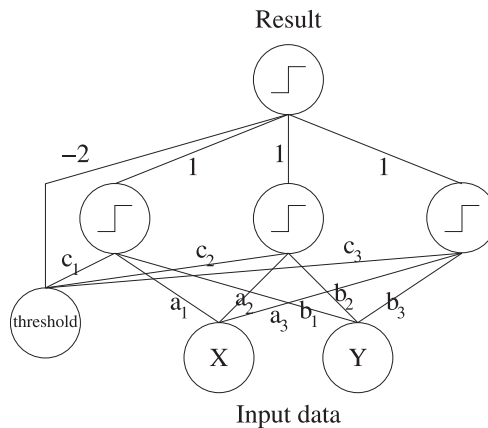


Рис. 3. Схема дискриминирующей функции (1)

Рассмотренная схема дискриминирующей функции представляет собой упрощённую модель трёхслойной нейронной сети прямого типа, в которой роль

нейронов играют пороговые дискриминаторы. Если заменить ПД гладкой функцией, например, сигмоидальной, нейронная сеть приобретёт очень важное свойство — способность обучаться. Такая сеть даёт возможность проводить классификацию образов больших размерностей, что практически невозможно реализовать традиционными методами.

Настройка сети (вычисление весов связей между нейронами и порогов) на решение конкретной задачи реализуется путём обучения, для чего чаще всего используется алгоритм обратного распространения ошибки [9]. Эта задача решается путём минимизации функционала ошибок E , представляющего собой сумму разностей между выходными сигналами сети и целевыми значениями:

$$E = \frac{1}{2} \sum_p \sum_i (\vec{y}_i^{(p)} - \vec{t}_i^{(p)})^2,$$

где p — индекс анализируемого образца, т.е. индекс события из набора данных, предназначенного для обучения сети, i — индекс выходного нейрона, $\vec{y}_i^{(p)}$ — выходной сигнал, получаемый с i -го нейрона для p -го образца, а $\vec{t}_i^{(p)}$ — целевое значение.

Процедура обучения нейронной сети повторяется до тех пор, пока значение выходного сигнала не приблизится к целевому значению. Обычно необходимо несколько проходов по всему набору данных, предназначенных для обучения, прежде, чем минимизируется функционал ошибок и будет получен оптимальный набор весов и порогов. По завершению этого процесса веса и пороги фиксируются, а качество обучения сети оценивается на основе тестового набора данных.

3. Выбор структуры сети

Задача идентификации электронов/пионов с помощью n -слойного TRD состоит в следующем: имея набор потерь энергии в n слоях TRD, нужно определить, какая частица, электрон или пион, была зарегистрирована детектором.

В качестве входной информации для сетей использовались выборки, составленные на основе потерь энергии в TRD пионами или электронами с импульсами от 1 ГэВ/с до 13 ГэВ/с. Энергетические потери получались путём Монте-Карло моделирования прохождения частиц через n -слойный TRD с помощью пакета GEANT [10] в среде CBM ROOT.

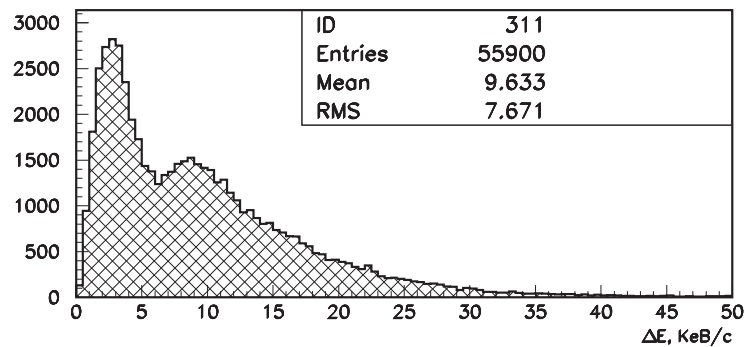
На рис. 4 приведены распределения потерь энергии электронов (а), включая потери на переходное излучение, и пионов (б) в первом слое-поглотителе детектора TRD для $p = 1,5$ ГэВ/с. Распределения потерь энергии в остальных слоях TRD и при других импульсах имеют аналогичный характер.

Исследовались две возможные причины различий в результатах работы рассматриваемых сетей:

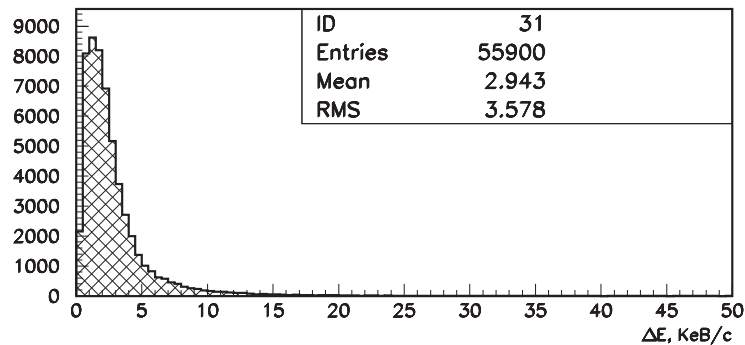
- 1) разные методы и алгоритмы минимизации функционала ошибок (смотри работу [3]);
- 2) различия в структуре сетей.

Проведённый анализ показал, что используемые на этапе обучения МСП методы минимизации функционала ошибок (алгоритмы Manhattan в JETNET и BFGS [11] в ROOT) не могут быть причиной заметных различий в результатах работы сетей. Вместе с тем в ходе этого анализа было обнаружено, что МСП из ROOT-а всегда даёт различающиеся результаты, т.е. сеть не воспроизводит полученные ранее результаты. Это приводит к тому, что нельзя зафиксировать порог для сигнала на выходе сети, по которому будет проводиться селекция электронов и пионов. Для того, чтобы исключить данное неудобство в работе сети, необходимо в списке параметров «option» функции Train задавать «+» [12].

В большинстве случаев параметры нейронных сетей с фиксированной архитектурой (у нас это МСП) настраиваются для решения конкретной задачи, и оптимальный вариант получается на основе эмпирического подбора. Однако, как



(a)



(б)

Рис. 4. Распределения потерь энергии электронов (а), включая потери на переходное излучение, и пионов (б) в первом слое-поглотителе детектора TRD для $p = 1,5$ ГэВ/с

оказалось в случае нашей задачи, даже небольшие различия в структуре сетей могут приводить к заметным различиям в результатах их работы.

При выборе структуры многослойного перцептрона необходимо задать:

- 1) количество слоёв МСП;
- 2) количество нейронов в каждом слое.

С количеством слоёв всё достаточно просто, так как для подавляющего большинства задач (наша задача относится к этому большинству), решаемых с помощью МСП, достаточно трёх слоёв: входной, скрытый и выходной.

Таким образом, нужно определиться с количеством нейронов в каждом слое. В обоих сетях число входных нейронов задавалось равным 12 (в соответствии с количеством слоёв в TRD). Для идентификации частицы использовался один выходной нейрон. При выборе числа нейронов в скрытом слое мы исходили из минимальной конфигурации сети [13], обеспечивающей оптимальный результат, а именно, максимальный уровень подавления пионов при минимальной потере электронов.

В качестве критерия для определения числа нейронов в скрытом слое использовалось распределение ошибок — разницы между целевым значением (для пионных событий целевое значение задавалось равным -1 , а для электронных — $+1$) и выходным сигналом сети. Распределение ошибок должно было удовлетворять следующим условиям:

- быть симметричным относительно нулевого среднего,
- иметь минимальное среднеквадратичное отклонение.

На рис. 5 показаны распределения ошибок на стадии обучения (а) и тестирования (в) МСП; правые гистограммы показывают распределения значений сигнала

на выходе нейронной сети, полученные на стадии её обучения (б) и тестирования (г) для $p = 1,5$ ГэВ/с.

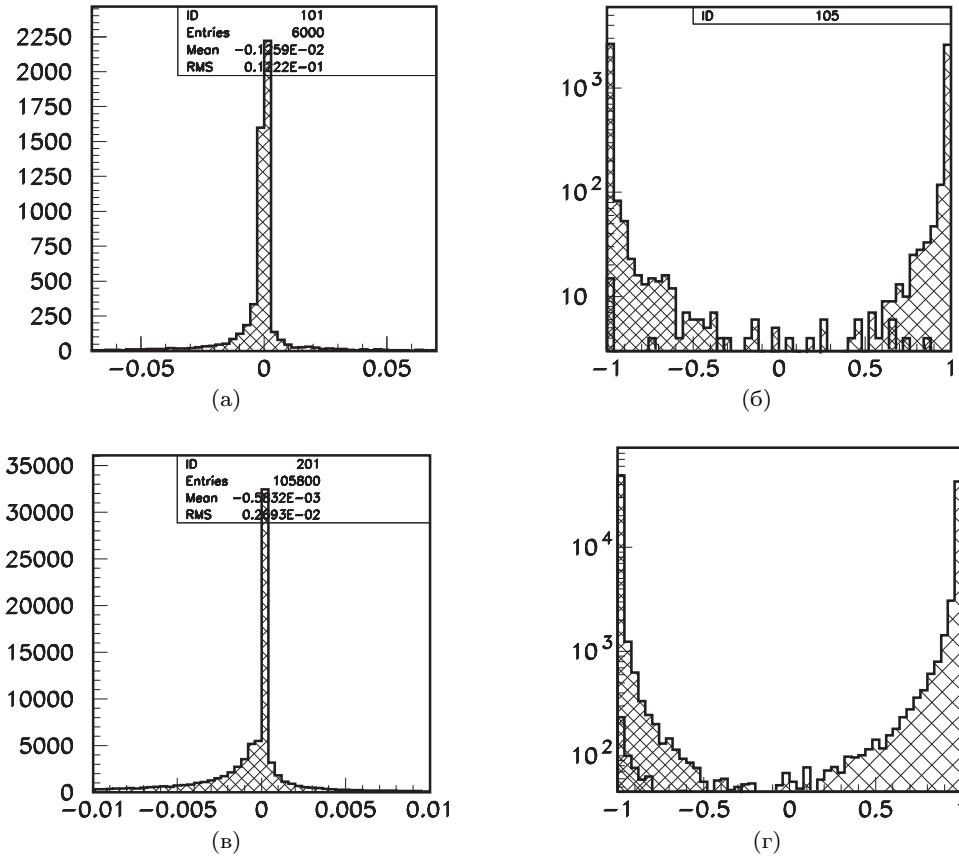


Рис. 5. Распределения ошибок на стадии обучения (а) и тестирования (в) МСП; распределения значений сигнала на выходе нейронной сети, полученные на стадии её обучения (б) и тестирования (г) для $p = 1,5$ ГэВ/с

4. Подготовка входных данных для МСП

Если в качестве входных данных для сети использовать выборки, составленные из потерь энергии электронов и пионов, то процесс обучения МСП проходит очень медленно, присутствуют большие колебания (относительно тренда) в эффективности распознавания частиц сетью.

В этой связи к исходным данным была применена следующая процедура преобразования (см. [3,14]):

$$\lambda_i = \frac{\Delta E_i - \Delta E_{mp}^i}{\xi_i} - 0,225, \quad i = 1, 2, \dots, n, \quad (2)$$

где ΔE_i — величина потери энергии зарегистрированной частицы (пиона или электрона) в i -м поглотителе TRD, ΔE_{mp}^i — величина наиболее вероятной потери энергии пиона в этом слое, $\xi_i = \frac{1}{4,02}$ FWHM; здесь FWHM — полная ширина на половине высоты для распределения потерь энергии пионов в i -м слое.

Для того, чтобы вычислить величины ΔE_{mp}^i и ξ_i , распределение потерь энергии пионов в i -м поглотителе (на данный момент все слои TRD имеют одинаковую

структуру) аппроксимировалось функцией плотности логнормального распределения (см. рис. 6)

$$f(x) = \frac{A}{\sqrt{2\pi\sigma x}} \exp^{-\frac{1}{2\sigma^2}(\ln x - \mu)^2}, \quad (3)$$

где σ — дисперсия, μ — среднее значение, а A — нормирующий множитель [15].

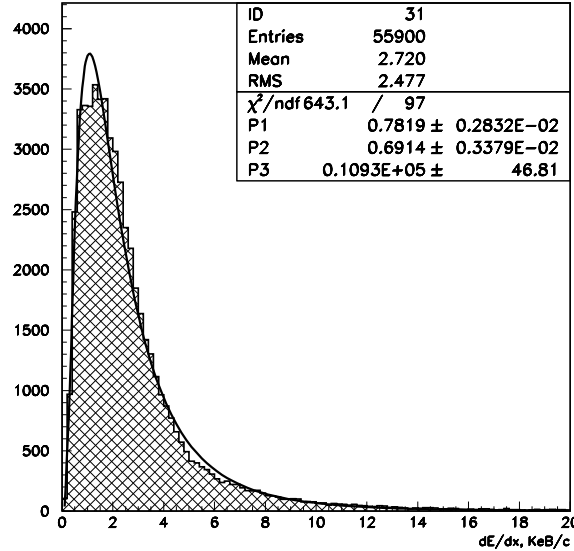


Рис. 6. Аппроксимация распределения потерь энергии пионов в первом поглотителе TRD с помощью функции плотности логнормального распределения

Элементы выборки λ_i упорядочивались по величине и использовались для формирования новой выборки из значений функции распределения Ландау $\phi(\lambda)$, вычисляемых с помощью функции DSTLAN (из библиотеки CERNLIB [16]). Подготовленная таким образом выборка величин $\phi(\lambda_j)$ ($j = 1, \dots, n$) использовалась в качестве образца, подаваемого на вход нейронной сети.

Применение «эффективных» переменных позволяет получить надёжный уровень идентификации электронов/пионов сетью после минимального количества эпох обучения. При этом практически отсутствуют колебания относительно тренда. Кроме того, довольно быстро достигается необходимый уровень подавления пионов при минимальной потере электронов (смотри поведение верхней кривой на рис. 7 и 8).

Следует отметить, что в случае использования исходных данных, несмотря на большое количество эпох обучения, не удаётся достичь приемлемого уровня распознавания частиц (смотри поведение нижней кривой на рис. 7 и 8).

5. Вычисление параметров преобразования

На рис. 9 и 10 приведены зависимости величин $\Delta E_{mp}(p)$ и $\xi(p)$, соответственно, от импульса пиона и их аппроксимация полиномом третьей степени: $f(p) = ap^3 + bp^2 + cp + d$.

В результате этой аппроксимации получены формулы для вычисления наиболее вероятной потери энергии ΔE_{mp}

$$\Delta E_{mp}(p) = 0,0005073p^3 - 0,01493p^2 + 0,1619p + 0,8978 \quad (4)$$

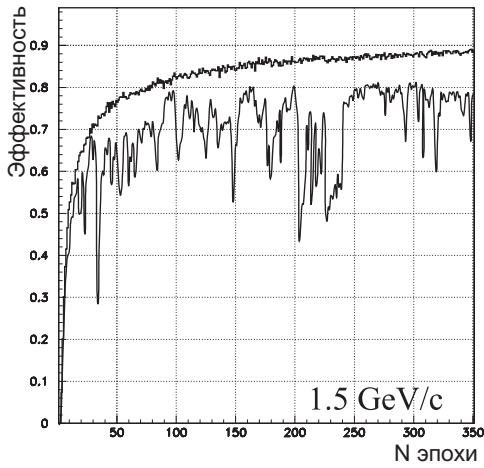


Рис. 7. Эффективности идентификации электронов/пионов с помощью МСП для исходных (нижняя кривая) и преобразованных (верхняя кривая) данных для $p = 1, 5$ ГэВ/с

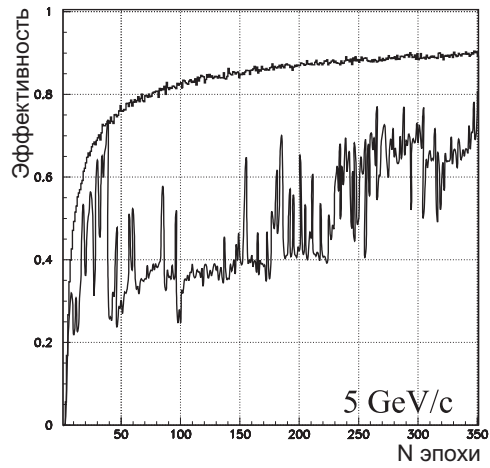


Рис. 8. Эффективности идентификации электронов/пионов с помощью МСП для исходных (нижняя кривая) и преобразованных (верхняя кривая) данных для $p = 5$ ГэВ/с

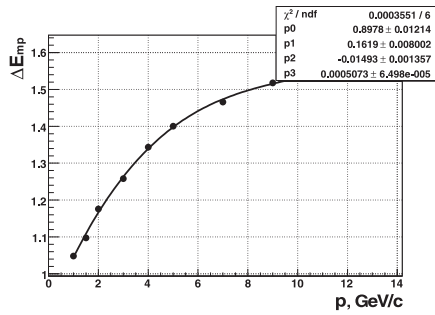


Рис. 9. Зависимость ΔE_{mp} от импульса пиона p (в ГэВ/с) и её аппроксимация полиномом третьей степени

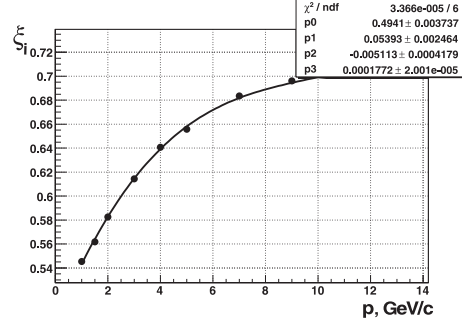


Рис. 10. Зависимость $\xi(p)$ от импульса пиона p (в ГэВ/с) и её аппроксимация полиномом третьей степени

и величины ξ

$$\xi(p) = 0,0001772p^3 - 0,005113p^2 + 0,05393p + 0,4941 \quad (5)$$

в зависимости от импульса пиона.

6. Обсуждение результатов

Для оценки эффективности идентификации электронов и подавления пионов с помощью МСП обе сети обучались и тестировались для каждого импульса в отдельности (см. табл. 1), используя соответствующие параметры преобразования (4) и (5). Под коэффициентом подавления пионов мы понимаем отношение полного числа проанализированных пионных выборок к тому числу пионных выборок, которые были идентифицированы как электроны.

В табл. 1 приведены зависимости коэффициентов подавления пионов от импульса при условии потери 10% электронов, полученные с помощью МСП из разных пакетов.

Таблица 1

Зависимости коэффициентов подавления пионов от импульса при условии потери 10 % электронов, полученные для разных МСП

p , ГэВ/с	1	1,5	2	3	4	5	7	9	11	13
JETNET	1857	1392	1837	1378	1713	1400	1317	1045	1089	897
ROOT	1219	814	1400	1112	1446	730	1054	610	882	654

Из табл. 1 видно, что в случае правильно выбранной структуры МСП обе нейронные сети обеспечивают высокий уровень подавления пионов и дают сопоставимые результаты.

7. Заключение

Проведено сравнение эффективностей работы многослойных перцептронов, взятых из двух разных пакетов (JETNET и ROOT), в задаче идентификации электронов и пионов с помощью детектора TRD в эксперименте CBM. Показано, что для правильно выбранной структуры МСП обе сети дают сопоставимые результаты. Даны рекомендации по выбору оптимальной конфигурации МСП. Использование сетей из разных пакетов даёт возможность проверки результатов, получаемых с помощью МСП. Тем самым обеспечивается надёжность физических результатов.

В широком импульсном интервале анализируемых частиц продемонстрировано, что применение «эффективных» переменных позволяет получить высокий уровень подавления пионов и обеспечивает надёжную идентификацию электронов сетью уже после минимального количества эпох обучения. При этом практически отсутствуют колебания относительно тренда. Получены аналитические формулы для вычисления параметров преобразования (необходимых для перехода к новым переменным).

Литература

1. Letter of Intent for the Compressed Baryonic Matter experiment. — <http://www.gsi.de/documents/DOC-2004-Jan-116-2.pdf>.
2. The CBM Collaboration. CBM Compressed Baryonic Matter Experiment. Technical Status Report: Techrep / GSI. — Darmstadt, 2005. — <http://www.gsi.de/onTEAM/dokumente/public/DOC-2005-Feb-447e.html>.
3. Electron/Pion Identification in the CBM TRD using a Multilayer Perceptron / E. P. Akishina, T. P. Akishina, V. V. Ivanov et al // JINR Communication. — 2007. — No E10-2007-17.
4. Peterson C., Rögnvaldsson T., Lönnblad L., 1994. — JETNET 3.0 – A Versatile Artificial Neural Network Package. — 81(1994)185.
5. TMultiLayerPerceptron / ROOT Reference Guide. — <http://root.cern.ch/root/html/TMultiLayerPerceptron.html>.
6. The Compressed Baryonic Matter experiment. — <http://www.gsi.de/fair/experiments/CBM>.
7. The FAIR Simulation and Analysis Framework / D. Bertini, M. Al-Turany, I. Koenig, F. Uhlig // Journal of Physics: Conference Series, Proc. of CHEP'07, Inter. Conference on Computing in High Energy and Nuclear Physics. — Victoria, BC Canada: 2007.
8. Denby B. Tutorial on Neural Networks Applications in High Energy Physics: 1982 Perspective // New Computing Techniques in Physics Research II, Proc. of the Second International Workshop on Software Engineering, Artificial Intelligence and Expert System in High Energy Physics. / Ed. by D. Perret-Gallix. — La Londe-les-Maures, France: World Scientific, 1992.

9. *Rumelhart D. E., Hinton G. E., Williams R. J.* Learning Internal Representations by Error Propagation // Parallel Distributed Processing: Explorations in the Microstructure of Cognition: Foundations / Ed. by D. E. Rumelhart, J. L. McClelland. — The M.I.T. Press, 1986. — Vol. 1. — Pp. 318–362.
10. 1995. — GEANT — Detector Description and Simulation Tool.
11. BFGS method. — http://en.wikipedia.org/wiki/BFGS_method.
12. TMultiLayerPerceptron / ROOT Reference Guide. — <http://root.cern.ch/root/html/TMultiLayerPerceptron.html>.
13. *Игнатъев Н. А.* Выбор минимальной конфигурации нейронных сетей // Вычислительные технологии. — 2001. — Т. 6, № 1. — С. 23–28.
14. *Zrelov P. V., Ivanov V. V.* The relativistic charged particles identification method based on the goodness-of-fit ω_n^3 -criterion // Nucl. Instr. and Methods in Phys. Res. — 1991. — Vol. A310. — Pp. 623–630.
15. Statistical Methods in Experimental Physics / W. T. Eadie, D. Dryard, F. E. James et al. — Amsterdam-London: North-Holland Pub.Comp, 1971.
16. *Koelberg K. S.* — CERN Computer Centre Program Library. — G110.

UDC 004.93'1:519.237:51-72

On Electron and Pion Identification using a Multilayer Perceptron in the Transition Radiation Detector of the CBM Experiment

T. P. Akishina, O. Yu. Derenovskaya, V. V. Ivanov

*Laboratory of Information Technologies
Joint Institute for Nuclear Research
141980, Dubna, Moscow region, Russia*

The problem of pion-electron identification based on their energy losses in the transition radiation detector (TRD) is considered in the frame of the CBM experiment. For particles identification an artificial neural network (ANN) was used, a multilayer perceptron being realized in JETNET and ROOT packages. It is demonstrated that, in order to get correct and comparable results, it is important to define the network structure correctly. The recommendations for such a selection are given. In order to achieve an acceptable level of pions suppression, the energy losses need to be transformed to more “effective” variables.

Key words and phrases: multivariate analysis, pattern recognition, CBM experiment, transition radiation detector TRD, neural networks.