

ПРИМЕНЕНИЕ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ПОИСКА РОЖДЕНИЯ БОЗОНА ХИГГСА СОВМЕСТНО С ОДИНОЧНЫМ ТОП-КВАРКОМ

И. Р. Бойко^{a,1}, *Н. А. Гусейнов*^{a,2}, *И. В. Елецких*^{a,3},
А. Р. Диденко^{a,б,4}, *О. А. Доловова*^{a,5}, *А. Д. Тропина*^{a,в,6}

^a Объединенный институт ядерных исследований, Дубна

^б Московский государственный университет им. М. В. Ломоносова, Москва

^в Московский физико-технический институт

(национальный исследовательский университет), Долгопрудный, Россия

Рассмотрена возможность применения искусственных нейронных сетей для поиска редкого процесса $pp \rightarrow tHbq$ на Большом адронном коллайдере. Разработан эволюционный алгоритм, позволяющий построить нейронную сеть с оптимальным набором параметров. При помощи моделирования методом Монте-Карло проведена оценка значимости выделения сигнала из фона. Применение нейронных сетей значительно увеличивает значимость сигнала и позволяет обнаружить процесс $pp \rightarrow tHbq$ на основе данных завершенных сеансов Большого адронного коллайдера.

We consider a possibility to apply Artificial Neural Networks to study the rare Higgs boson production process $pp \rightarrow tHbq$ at the Large Hadron Collider. We have developed an evolutionary algorithm for optimization of the parameter set of the neural network. Signal significance is evaluated using Monte Carlo simulation. Application of the neural network increases the signal significance, allowing an observation of the signal process with the currently available LHC statistics.

PACS: 14.80.+Bn; 12.60.+i

ВВЕДЕНИЕ

В последнее время методы машинного обучения находят все более широкое применение при анализе экспериментальных данных в физике частиц. Усложнение экспериментальных установок приводит к необходимости совместного анализа большого

¹Е-mail: boyko@jinr.ru

²Е-mail: nguseynov@jinr.ru

³Е-mail: ivaneleckih@jinr.ru

⁴Е-mail: alisadidenko@jinr.ru

⁵Е-mail: kovo@jinr.ru

⁶Е-mail: atropina@jinr.ru

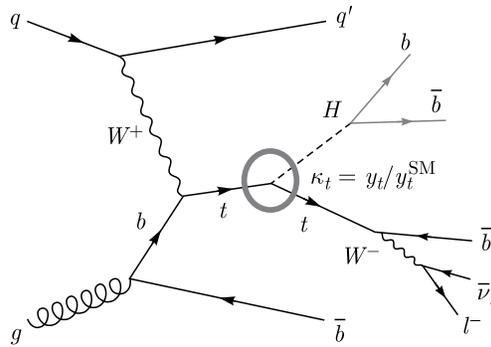


Рис. 1. Диаграмма Фейнмана для t -канала сигнального процесса $pp \rightarrow tH$ с распадом бозона Хиггса на пару $b\bar{b}$

количества наблюдаемых величин. При этом нередко каждая из индивидуальных величин обладает незначительной чувствительностью к физическим параметрам, изменение которых является целью эксперимента. Для извлечения из экспериментальных данных полной информации необходимо прибегать к мультивариативному анализу, который принимает во внимание как распределения измеряемых величин, так и корреляции между ними.

В настоящей работе рассматривается возможность применения методов машинного обучения (нейронных сетей) для поиска редкого процесса — рождения бозона Хиггса совместно с одиночным топ-кварком. На протон-протонных коллайдерах уравнение данной реакции можно записать как $pp \rightarrow tHbq$ в первом порядке теории возмущений. Диаграмма процесса приведена на рис. 1. Топ-кварк и бозон Хиггса (tH) рождаются совместно с b -кварком (b) и кварком-спектатором q (обычно u - или d -кварк). При этом b -кварк рождается с малым поперечным импульсом и, как правило, не регистрируется в эксперименте. В то же время кварк-спектатор регистрируется как адронная струя высокой энергии в области больших псевдобыстрот.

Настоящее исследование основано на моделировании методом Монте-Карло сигнала $pp \rightarrow tHbq$ (сокращенно tH) и основных фоновых процессов. Моделированные данные описаны в работе [1], где разделение сигнала и фона производилось методом последовательных обрезаний. В настоящей работе анализ тех же данных производится при помощи искусственных нейронных сетей (ИНС), что позволяет существенно улучшить выделение сигнала из фоновых процессов.

При анализе данных не проводилось детального моделирования экспериментальной установки. В то же время учитывались типичные характеристики экспериментов ATLAS и CMS на Большом адронном коллайдере (БАК): эффективность регистрации струй, энергетическое разрешение и др. (см. [1]). Рассматривались каналы распада $H \rightarrow b\bar{b}$, $t \rightarrow b\ell\nu_\ell$ ($\ell = e, \mu$), обеспечивающие относительно высокую статистику событий сигнала. Таким образом, события сигнала $pp \rightarrow tHbq$ характеризовались наличием в конечном состоянии заряженного лептона с большим поперечным импульсом, недостающего поперечного импульса от нейтрино, трех или четырех струй от b -кварков (b -струй), а также «передней струи» от кварка-спектатора, характе-

ризующейся большой полной энергией и большим абсолютным значением псевдо-быстроты.

Основным фоновым процессом являлось парное рождение топ-кварка и топ-антикварка $pp \rightarrow tt$ (здесь и далее не делается различия между частицами и античастицами, если это не приводит к недоразумению). Моделировался распад пары топ-кварков по полулептонному каналу $tt \rightarrow (b\nu_\ell)(bqq')$, где кварк-антикварковая пара qq' — это либо ud , либо sc . Таким образом, в конечном состоянии имеется два b -кварка, в то время как сигнал характеризуется наличием не менее чем трех b -кварков. Однако при моделировании учитывалась возможность ошибочной идентификации b -струи [2, 3], особенно при реконструкции струи от c -кварка. Такие ложно реконструированные b -струи составляют главную часть фона. В качестве отдельного фонового процесса выделяются события $pp \rightarrow ttbb$, где «недостающие» b -кварки рождаются благодаря квантовым поправкам к каналу парного рождения $pp \rightarrow tt$.

В качестве дополнительных фоновых процессов моделировались реакции $pp \rightarrow ttH$ и $pp \rightarrow tZ$. Все сигнальные и фоновые процессы моделировались генератором MadGraph 5 [4]. Для сигнала $pp \rightarrow tH$ было смоделировано два набора событий: согласно предсказаниям Стандартной модели (СМ) и согласно модели инвертированного взаимодействия топ-кварка (ИВТ). Модель ИВТ, предложенная в работах [5–8], является расширением Стандартной модели, в котором юкавовское взаимодействие топ-кварка y_t совпадает с предсказанием СМ по абсолютной величине, но противоположно ему по знаку. ИВТ предсказывает, что сечение процесса $pp \rightarrow tH$ более чем на порядок превышает предсказание СМ [9], что значительно улучшает чувствительность к обнаружению этого редкого канала.

В недавней электронной публикации эксперимента CMS [10] представлен совместный анализ каналов tH и ttH . Для канала tH установлен довольно слабый верхний предел сечения, приблизительно в 15 раз превышающий предсказание СМ. Невысокая точность измерений не позволила исключить предсказания модели ИВТ или определить знак константы юкавовского взаимодействия. Дальнейшее улучшение чувствительности может быть достигнуто с появлением новых экспериментальных данных или за счет совершенствования методов их анализа.

ПРЕДВАРИТЕЛЬНАЯ ПОДГОТОВКА ДАННЫХ

В табл. 1 суммируются сведения о моделированных событиях сигнала и фона. Число событий, ожидаемых на Большом адронном коллайдере, вычислялось для интегральной светимости 140 фб^{-1} , т. е. для статистики, набранной каждым из экспериментов ATLAS И CMS в последнем завершённом сеансе Run-2. Сечения указаны с учетом вероятностей распадов $H \rightarrow bb$ и $t \rightarrow b\nu_\ell$. В дальнейшем анализе использовались события, удовлетворяющие следующему предварительному отбору: реконструировано не менее трех b -струй и не менее одной «передней струи» с поперечным импульсом не менее $25 \text{ ГэВ}/c$ и полярным углом относительно оси пучков не более 20° . Количество сгенерированных событий в табл. 1 указано с учетом этих критериев отбора. Полное количество сгенерированных событий приведено в [1].

Таблица 1. Сечения промоделированных процессов и количество сгенерированных событий

| Параметр | Процесс | | | | | |
|--------------------------------|----------------|-----------------|-----------|-------------|------------|-------------|
| | <i>tH</i> (СМ) | <i>tH</i> (ИВТ) | <i>tt</i> | <i>ttbb</i> | <i>ttH</i> | <i>tZbq</i> |
| Сечение, фб | 10,8 | 97 | 306 000 | 3600 | 72 | 18 |
| Число сгенерированных событий | 12500 | 14057 | 13 994 | 7268 | 5282 | 3750 |
| Ожидаемое на БАК число событий | 206 | 3145 | 161 233 | 30027 | 1441 | 461 |

Возможность наблюдения сигнала при наличии значительного фона принято характеризовать величиной значимости *Z*, вычисляемой по формуле

$$Z = \frac{S}{\sqrt{S + B}}, \tag{1}$$

где *B* — количество событий фона, а *S* — количество событий сигнала. В качестве критерия открытия принято считать величину значимости, превышающую 5σ . Для статистики, указанной в табл. 1, значимость составляет $0,47\sigma$ для сигнала СМ и $7,1\sigma$ для сигнала ИВТ. Основной задачей данной работы являлось увеличение уровня значимости сигнала.

Среди кинематических и иных переменных, характеризующих события сигнала и фона, были выбраны 37 переменных, распределения которых обеспечивают наилучшее разделение сигнала и фона. Разделяющая способность переменной (SP, от Separation Power, см., например, [11]) вычисляется по формуле

$$SP = \frac{1}{2} \sum_i \frac{(s_i - b_i)^2}{(s_i + b_i)}, \tag{2}$$

где s_i (b_i) — это содержимое *i*-го канала гистограммы распределения сигнала (фона). Оба распределения должны быть отнормированы на единичную площадь, т.е. сумма всех каналов каждого распределения должна равняться единице. Разделяющая способность принимает значения от 0 до 100%, а ее приблизительный физический смысл — это доля статистики, которая позволила бы при полном отсутствии фона измерить сигнал с той же точностью, с которой сигнал фактически измеряется при наличии фона.

Список выбранных переменных и разделяющая способность каждой из них приведены в табл. 2. Распределения некоторых переменных приведены на рис. 2.

Как видно из рис. 2, различные переменные варьируются в сильно отличающихся диапазонах. В то же время желательно (хотя и не обязательно), чтобы входные переменные нейронной сети имели схожий диапазон изменения. В противном случае время работы (тренировки) нейронной сети может сильно увеличиться. Для решения этой проблемы существуют методы «нормализации» переменных. Нами использовалась процедура, в которой для каждой переменной производится преобразование, переводящее распределение событий фона по этой переменной в нормальное распределение Гаусса. То же самое преобразование применяется и к распределению сигнала по данной переменной. Значения переменных, полученные в результате нормализации, используются в качестве входных параметров нейронной сети. Процедура нормализации проиллюстрирована на рис. 3.

Таблица 2. Набор переменных, наиболее чувствительных к отбору событий сигнального процесса $pp \rightarrow tH$. Термин «объекты» означает совокупность всех струй заряженного лептона и нейтрино

| Название | Разделяющая способность | | Определение |
|--------------------------------|-------------------------|-------|---|
| | СМ | ИВТ | |
| M_{3J} | 12,35* | 6,71 | Инвариантная масса трех струй с наибольшим p_T |
| N_b | 10,71 | 9,41 | Количество струй, порожденных b -кварками |
| Sphresity _{all jets} | 9,23 | 3,72 | Мера равномерности распределения струй в пространстве |
| Sphresity _{1nu} | 9,02 | 3,05 | Мера равномерности распределения объектов в пространстве |
| Aplanarity _{all jets} | 8,92* | 5,79 | Мера отклонения струй от одной общей плоскости |
| $\Delta(\eta_t, H)$ | 8,54 | 0,66 | Разность псевдобыстрот топ-кварка и бозона Хиггса |
| Aplanarity _{1nu} | 8,37 | 3,86 | Мера отклонения объектов от общей плоскости |
| $\Delta(\eta_t, FWD)$ | 7,47 | 4,93* | Разность псевдобыстрот топ-кварка и передней струи |
| $M_{t, H}$ | 6,94* | 1,44* | Инвариантная масса топ-кварка и бозона Хиггса |
| M_H | 6,60 | 6,75 | Восстановленная масса бозона Хиггса |
| $f_{wmlnujet1}$ | 5,61 | 2,97 | Первый момент Фокса–Вольфрама, составленный из импульсов объектов |
| P_t^{FWD} | 5,48 | 3,45 | Поперечный импульс передней струи |
| $M_{H, FWD}$ | 5,45 | 11,17 | Инвариантная масса бозона Хиггса и передней струи |
| $M_{H, cen. jet.}$ | 5,44 | 3,00 | Инвариантная масса бозона Хиггса и центральной легкой струи |
| η_{FWD} | 5,11 | 6,08 | Псевдобыстрота передней струи |
| χ^{\min} | 4,28 | 2,78 | Качество (критерий χ^2) определения масс бозона Хиггса и топ-кварка |
| M_t | 3,61 | 2,31 | Восстановленная масса t -кварка |
| Q_{lep} | 3,32 | 2,80 | Заряд лептона с наибольшим p_T |
| $\Delta(R_{qqW})$ | 3,06 | 3,99* | Угол между струями от адронного распада W -бозона |
| N_{nonb} | 2,79 | 3,97 | Количество струй, порожденных кварками, отличным от b -кварка |
| f_{wm2} | 2,70* | 0,96* | Второй момент Фокса–Вольфрама, составленный из импульсов струй |
| RapGap_maxptb | 2,08 | 6,82 | Разность псевдобыстрот передней струи и b -струи с наибольшим p_T |
| RapGap_closestb | 1,94 | 8,16 | Разность псевдобыстрот передней струи и ближайшей к ней b -струи |
| $P_{nonb_max}^t$ | 1,91 | 0,92 | Наибольший поперечный импульс среди легких струй |
| W_T_m | 1,83* | 1,29 | Поперечная масса всех струй |
| M_{FWD} | 1,79* | 0,96 | Инвариантная масса передней струи и топ-кварка |
| η_{lep} | 1,73 | 2,43* | Псевдобыстрота лептона с наибольшим p_T |
| E_{b2} | 1,62 | 0,69 | Энергия третьей по поперечному импульсу b -струи |
| $\Delta(\phi_{t, H})$ | 1,49* | 0,36 | Разность азимутальных углов топ-кварка и бозона Хиггса |
| HT_alljets | 1,43 | 2,32 | Алгебраическая сумма поперечных импульсов всех струй |
| $\Delta(\eta_H, FWD)$ | 1,41* | 5,70 | Разность псевдобыстрот бозона Хиггса и передней струи |
| P_H^t | 1,15 | 4,33 | Восстановленный поперечный импульс бозона Хиггса |
| P_{b0}^t | 1,07 | 3,57 | Поперечный импульс b -струи с наибольшим p_T |
| P_{b1}^t | 0,86 | 2,22* | Поперечный импульс второй по p_T b -струи |
| η_H | 0,74 | 2,00* | Восстановленная псевдобыстрота бозона Хиггса |
| η_{b1} | 0,61 | 1,29 | Псевдобыстрота второй по p_T b -струи |
| M_{b0} | 0,38 | 1,49* | Инвариантная масса передней струи и b -струи с наибольшим p_T |

*Исключена эволюционным алгоритмом для данной модели.

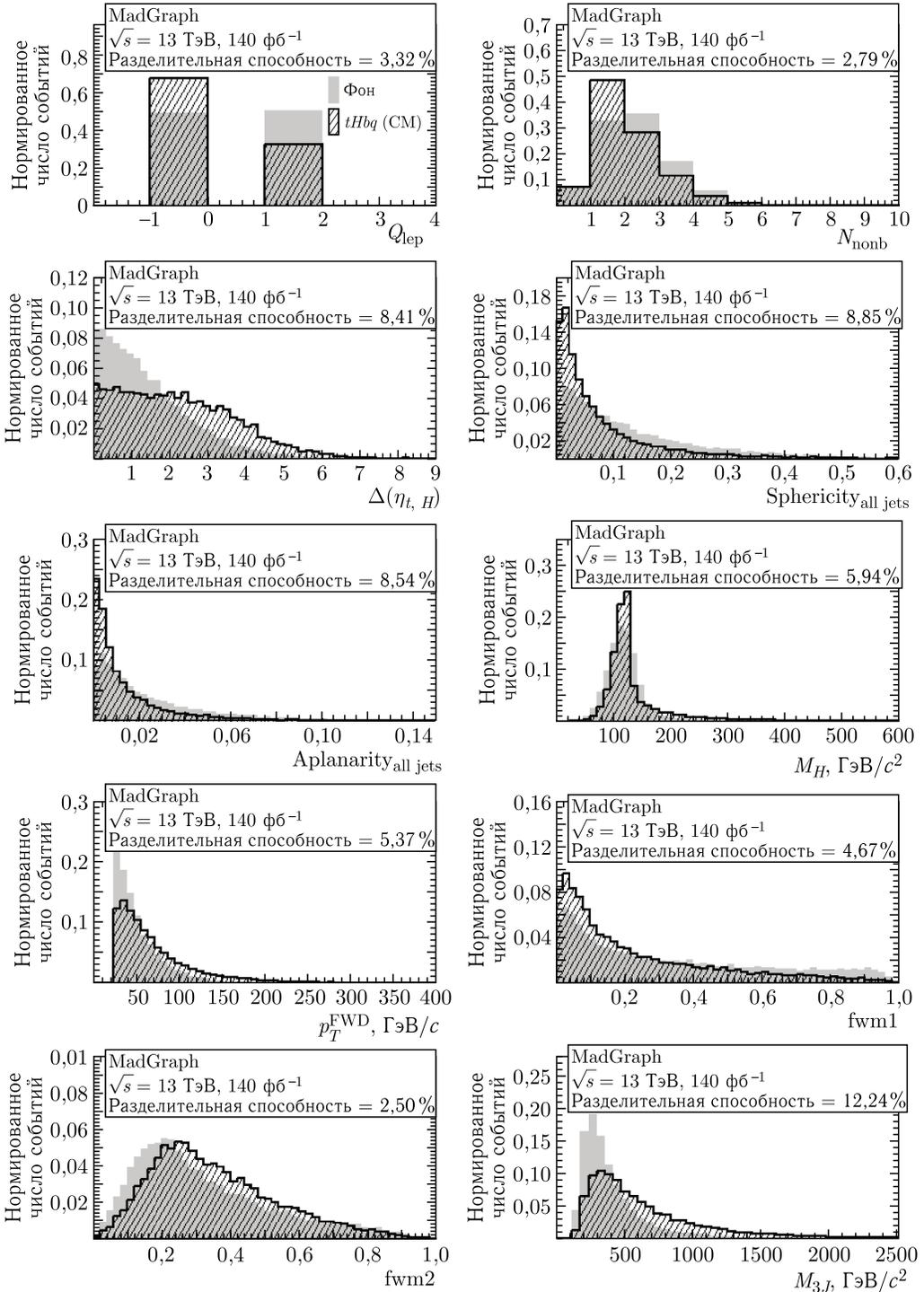


Рис. 2. Распределение некоторых входных кинематических переменных. Фон включает процессы tt , ttH , $tZbq$

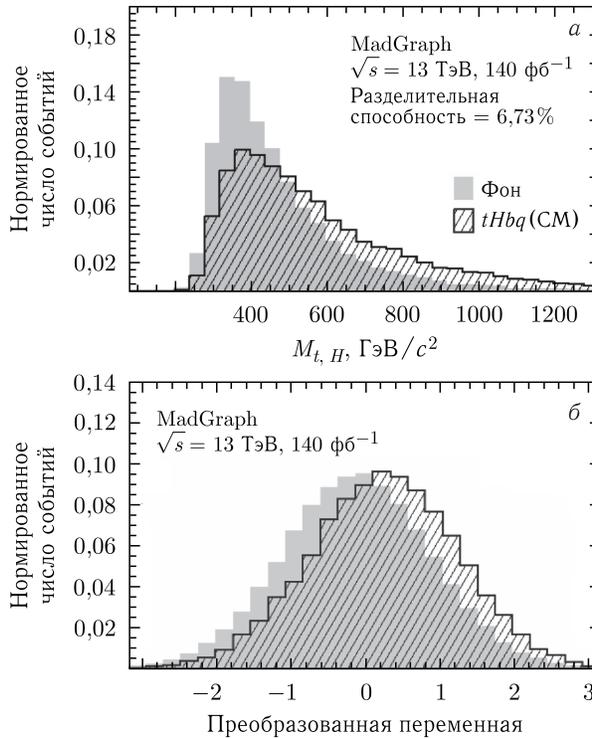


Рис. 3. а) Нормированное распределение инвариантной массы t -кварка и бозона Хиггса; б) нормированное распределение $M_{t,H}$ после преобразования к распределению Гаусса. Фон включает процессы tt , ttH , $tZbq$

ПАРАМЕТРЫ И ГИПЕРПАРАМЕТРЫ НЕЙРОННЫХ СЕТЕЙ

При обучении искусственной нейронной сети (ИНС) происходит оптимизация большого количества (тысяч) параметров ИНС — весов и сдвигов, описывающих связи между нейронами. Помимо этих «явных» параметров, нейронная сеть также характеризуется набором неявных параметров, или «гиперпараметров», от значений которых может существенно зависеть эффективность ее работы:

- а) *структура сети* — число скрытых слоев и число нейронов в каждом из них;
- б) *функции активации нейронов* (\tanh , relu , sigmoid , exponential) [11, с. 95];
- в) *алгоритм обучения* (SGD, Adam, RMSprop) и его собственные параметры, например, число параметров ИНС, оптимизируемых за одну итерацию алгоритма обучения, величина обратной связи между параметрами ИНС и откликом сети в процессе обучения и т. п.;
- г) *набор входных переменных*. Задача обучения сети существенно усложняется при увеличении числа входных переменных. Помимо роста времени обучения, существенно увеличивается вероятность попадания в локальный минимум в пространстве параметров ИНС. Помимо этого некоторые кинематические переменные могут иметь высокую корреляцию друг с другом. Использование в ИНС таких скоррелированных

переменных зачастую дает эффект, обратный желаемому, — эффективность работы сети падает, несмотря на добавление дополнительной информации. Таким образом, выбор оптимального набора входных переменных сети, как правило, является важной и нетривиальной задачей.

Все общепринятые алгоритмы обучения нейронных сетей предназначены для оптимизации только «явных» параметров ИВН, описывающих связи между нейронами.

При этом оптимизация гиперпараметров осуществляется либо вручную пользователем (как правило, на основе собственного опыта и интуиции), либо специальными отдельными алгоритмами. Оптимизация гиперпараметров может существенно влиять на результат классификации экспериментальных событий, однако ее эффективность не всегда высока.

Нами был разработан алгоритм автоматической оптимизации гиперпараметров искусственной нейронной сети. Задача оптимизации гиперпараметров ИНС сводится к задаче поиска глобального минимума в многомерном пространстве гиперпараметров. Оптимизируемой величиной (или иначе «целевой функцией») здесь может быть значение ROC-интеграла, статистическая значимость сигнала над фоном или любой иной критерий качества разделения сигнала и фона.

ОПТИМИЗАЦИЯ ГИПЕРПАРАМЕТРОВ ИНС

Особенность пространства гиперпараметров заключается в том, что часть его измерений дискретна, часть непрерывна или квазинепрерывна. По многим из них пространство недифференцируемо и представляет собой так называемые «скалы Хэмминга». Иными словами, соседние значения дискретных (или квазинепрерывных) гиперпараметров зачастую соответствуют огромной разнице в значении целевой функции. В таких пространствах методы оптимизации, в той или иной степени зависящие от предсказания профиля пространства параметров (градиентный спуск или байесовская оптимизация, как в работе [10] и т. п.), зачастую страдают от попадания в локальный минимум, далекий от глобального. С другой стороны, полностью стохастические алгоритмы, использующие случайный поиск или методы Монте-Карло, требуют слишком большого количества вычислений целевой функции для нахождения глобального или достаточно глубокого локального минимума. Более того, это время растет очень быстро с ростом числа измерений пространства параметров. Помимо этого задача оптимизации гиперпараметров является примером нерегулярной задачи, так как значение целевой функции в заданной точке не детерминировано и само по себе зависит от результата работы алгоритма обучения, как правило использующего стохастические методы и поэтому не обладающего повторяемостью результата.

Одним из признанных методов, демонстрирующих высокую эффективность в сложных или нерегулярных задачах оптимизации, является применение эволюционных алгоритмов (см., например, [12]). Суть эволюционного алгоритма заключается в создании вначале случайного (или определенного каким-либо иным алгоритмом) набора решений (точек в пространстве параметров), имеющего название популяции. На первом шаге в популяции отбирается одно или несколько лучших решений, а оставшиеся решения из популяции удаляются. Вторым шагом производится «размножение» лучших решений — они «клонировются» или «скрещиваются». Наконец, на третьем шаге

новые решения получают (с определенной вероятностью) случайные приращения оптимизируемых параметров (так называемые мутации). Эти новые решения называются потомками и занимают освободившееся место в популяции. Цикл из таких трех шагов называется поколением эволюционного алгоритма. Циклически повторяя отбор, копирование и мутации, алгоритм работает в течение определенного числа поколений, пока не будет выполнено условие остановки. Таким условием может быть нахождение достаточно хорошего значения целевой функции, снижение скорости оптимизации ниже определенного порога, прохождение заданного числа поколений и др.

При разумном выборе степени наследственности и алгоритмов изменчивости эволюционные алгоритмы эффективно справляются с оптимизацией в сложных пространствах.

Эволюционный алгоритм такого типа был применен нами для поиска оптимального набора гиперпараметров нейронной сети (перечисленных в предыдущем разделе) в задаче поиска сигнала рождения бозона Хиггса в ассоциации с топ-кварком.

Создается популяция из 30 случайных наборов гиперпараметров. Для каждого набора гиперпараметров проводится обучение нейронной сети, предназначенной для разделения сигнала и фона. Обученная нейронная сеть присваивает каждому событию значение отклика сети N_{out} , находящееся в диапазоне от 0 до 1. Близкие к нулю значения характерны для фоновых событий, близкие к единице — для сигнала. Для всей статистики моделированных событий вычисляется целевая функция, в качестве которой в данной работе использовалось значение ROC-интеграла (см. ниже). Выбирается набор гиперпараметров, дающий наилучшую целевую функцию, а остальные наборы удаляются из популяции. Лучшее решение копируется 29 раз, и в каждую копию вносятся случайные вариации гиперпараметров (мутации).

При этом вероятности мутаций разных параметров существенно различны в соответствии с ожидаемым потенциальным эффектом от изменения данного гиперпараметра на целевую функцию. К примеру, небольшое изменение числа нейронов в конкретном слое редко оказывает значительное влияние на результат. Вероятность таких мутаций может быть сделана большой. В то же время изменение алгоритма обучения сети или набора входных переменных влияет на результат сильнее, вероятность мутаций этих параметров выбирается небольшой. Вероятность выключения переменных из входного набора меньше вероятности возвращения переменных во входной набор, поскольку больший набор переменных, как правило, вероятнее даст лучшее решение. За счет правильного выбора вероятностей мутаций для каждого набора «важных» параметров (например, набора входных переменных) алгоритм успевает оптимизировать оставшиеся «менее важные» параметры (например, число слоев и нейронов). Это в целом позволяет ускорить работу алгоритма.

РЕЗУЛЬТАТЫ

Описанный выше эволюционный алгоритм был применен для оптимизации гиперпараметров искусственной нейронной сети, предназначенной для отделения сигнала $pp \rightarrow tH$ от фоновых событий. Оптимизация проводилась отдельно для сигнала Стандартной модели и сигнала модели ИВТ. Оптимальные наборы гиперпараметров приведены в табл. 3.

Таблица 3. Набор оптимальных гиперпараметров для сигнала Стандартной модели и сигнала ИВТ

| Название | Сигнал | |
|---|------------------|---------------|
| | СМ | ИВТ |
| Входной слой | 29 переменных | 29 переменных |
| Скрытый слой 1 | 1878 (relu) | 1853 (relu) |
| Скрытый слой 2 | 264 (relu) | 208 (relu) |
| Скрытый слой 3 | 52 (exponential) | 55 (sigmoid) |
| Скрытый слой 4 | 74 (exponential) | 67 (softmax) |
| Общее число параметров обучения | 570 248 | 456 605 |
| Количество параметров, обучаемых за раз | 87 | 96 |
| Алгоритм обучения | RMSprop | RMSprop |

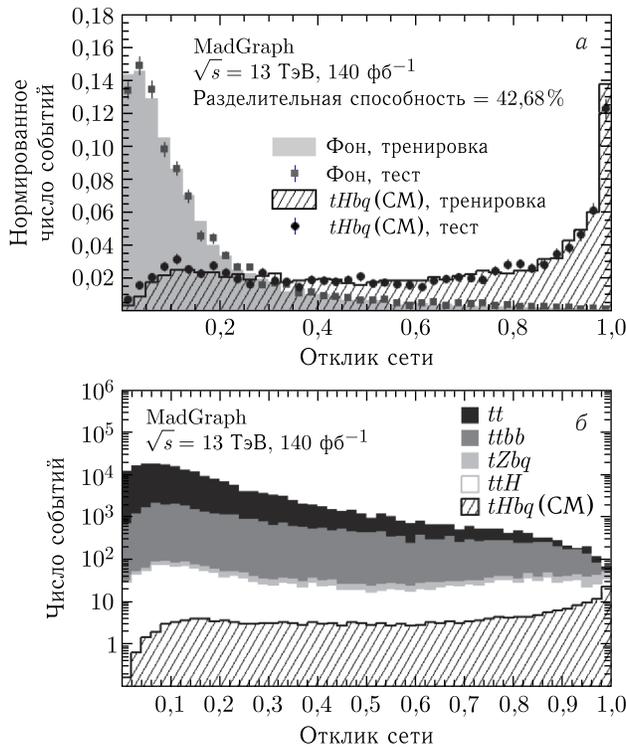


Рис. 4. а) Нормированное распределение отклика сети для тренировочных и тестовых выборок, где фон включает процессы $tt, ttH, tZbq$; б) отклик сети N_{out} с гиперпараметрами, оптимальными для сигнала СМ

Моделированные данные были обработаны нейронными сетями с описанными выше оптимальными наборами гиперпараметров. На рис. 4 показано распределение сигнала и фона по значениям отклика сети для сигнала Стандартной модели. Как и ожидается, фон концентрируется вблизи нуля, а сигнал — вблизи единицы. Тренировка нейронной сети производилась с использованием случайной выборки, включающей

3/4 от всей имеющейся статистики. На рис. 4, а показано сравнение этой выборки с «тестовой статистикой», т. е. с оставшейся 1/4 событий, не использовавшихся для тренировки. Хорошее согласие между распределениями показывает отсутствие перетренировки. Разделяющая способность распределения отклика сети составляет 42,7%. Для сигнала ИВТ разделяющая способность отклика сети составляет 44,3%.

На рис. 5 показаны ROC-кривые для нейронных сетей с оптимизированными гиперпараметрами. Кривые показывают соотношение между эффективностью отбора сигнала и долей отбрасываемых фоновых событий при различных значениях параметра обрезания по значению отклика сети. Интеграл кривой (ROC-интеграл) составляет 0,90 для СМ и 0,88 для ИВТ.

На рис. 6 проиллюстрирована достигнутая значимость сигнала. Для каждого из значений отклика сети N_{out} (показаны по горизонтальной оси) значимость сигнала вычислялась для событий со значениями выходного нейрона, превышающими N_{out} . Для уменьшения влияния статистических погрешностей оптимальная значимость определялась с учетом дополнительного условия: после обрезания по величине N_{out} должно быть отобрано не менее 25% событий сигнала. Для сигнала СМ оп-

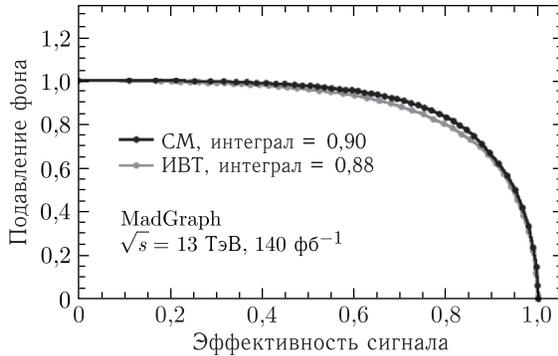


Рис. 5. Кривые эффективности для сигнала СМ и ИВТ

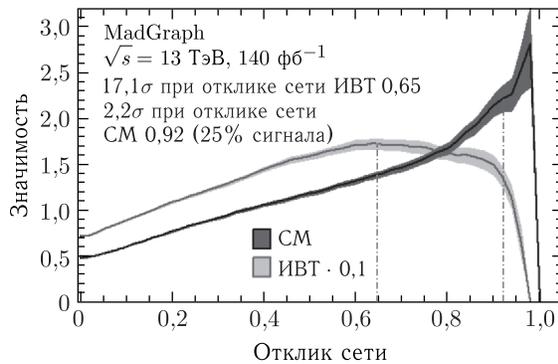


Рис. 6. Значимость сигналов СМ и ИВТ (уменьшена в 10 раз). Серые полосы вокруг кривых показывают статистическую погрешность, вызванную ограниченной статистикой моделированных событий. Вертикальные линии показывают оптимальные параметры обрезания

тимальная значимость достигается при значении $N_{\text{out}} = 0,92$ (соответствует отбору 25 % сигнала) и составляет $2,2\sigma$. Это в 4,7 раза превышает значимость сигнала без применения ИНС (см. табл. 1), что эквивалентно увеличению экспериментальной статистики более чем в 20 раз. Для сигнала ИВТ оптимальная значимость ($17,1\sigma$) достигается при значении параметра обрезания $N_{\text{out}} = 0,65$.

ЗАКЛЮЧЕНИЕ

Возможность обнаружения редкого канала рождения бозона Хиггса $pp \rightarrow tHbq$ проанализирована на основе данных, смоделированных методом Монте-Карло, с применением искусственных нейронных сетей. Разработан эволюционный алгоритм, позволяющий автоматически оптимизировать набор гиперпараметров нейронной сети. Применение алгоритма позволяет повысить значимость сигнала более чем в 4 раза (сигнал СМ) или более чем в 2 раза (сигнал ИВТ).

Значения значимости сигнала, приведенные для светимости БАК, 140 фб^{-1} , учитывают только статистические погрешности измерений. Систематические погрешности экспериментов в настоящей работе не рассматривались. Если предположить, что влияние систематических погрешностей уменьшает значимость вдвое, то ожидаемая значимость сигнала Стандартной модели составляет порядка 1σ . Таким образом, по результатам сеанса Run-2 эксперименты БАК не смогут надежно выделить сигнал tH , предсказываемый Стандартной моделью. В то же время значимость сигнала модели ИВТ может составить порядка 8σ даже с учетом систематических погрешностей. Следовательно, применение предложенного алгоритма оптимизации позволило бы обнаружить или исключить сигнал ИВТ на основе данных завершеного сеанса Run-2.

СПИСОК ЛИТЕРАТУРЫ

1. *Boyko I., Huseynov N., Koval O.* Monte-Carlo Study of Associated Higgs Boson Production with a Single Top-Quark // *Yad. Fiz.* 2022. V. 85, No. 2. P. 120–128.
2. *ATLAS Collab.* // *J. Instrum.* 2016. V. 11. P. P04008.
3. *CMS Collab.* // *J. Instrum.* 2018. V. 13. P. P05011.
4. *Alwall J. et al.* The Automated Computation of Tree-Level and Next-to-Leading Order Differential Cross Sections, and Their Matching to Parton Shower Simulations. arXiv:1405.0301. 2014.
5. *Biswas S., Gabrielli E., Mele B.* // *JHEP.* 2013. V. 01. P. 088.
6. *Farina M. et al.* // *Ibid.* V. 05. P. 022.
7. *Agrawal P., Mitra S., Shivaji A.* // *Ibid.* V. 12. P. 077.
8. *Biswas S. et al.* // *Ibid.* V. 07. P. 073.
9. LHC Higgs Cross Section Working Group. Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector. CERN. Geneva, 2016.
10. *CMS Collab.* Measurement of the ttH and tH Production Rates in the $H \rightarrow bb$ Decay Channel. CMS-PAS-HIG-19-011; <http://cds.cern.ch/record/2868175>. 2023.
11. *Hoecker A. et al.* TMVA — Toolkit for Multivariate Data Analysis. 2009; arXiv:physics/0703039.
12. *Tamilselvi S.* Introduction to Evolutionary Algorithms. IntechOpen, 2022.

Получено 15 декабря 2023 г.