

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ  
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ  
«МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ имени  
М.В.ЛОМОНОСОВА»

ФИЗИЧЕСКИЙ ФАКУЛЬТЕТ

КАФЕДРА ФИЗИКИ ЭЛЕМЕНТАРНЫХ ЧАСТИЦ

**ПРИМЕНЕНИЕ МЕТОДОВ МАШИННОГО  
ОБУЧЕНИЯ  
ДЛЯ ОЦЕНКИ ПОТОКОВ ЧЕРЕНКОВСКОГО  
ИЗЛУЧЕНИЯ В ЭКСПЕРИМЕНТЕ БАКАЛ-GVD**

Студент: Жаворонкова В. И.  
Научный руководитель: Наумов Д. В.  
Научный консультант: Аллахвердян В. А.

Группа: 218

Москва  
2024

# Оглавление

<b>1</b>	<b>Введение</b>	<b>2</b>
<b>2</b>	<b>Теоретическая часть</b>	<b>2</b>
2.1	Кинетические уравнения . . . . .	2
2.2	Уравнение Больцмана . . . . .	3
2.3	Уравнение переноса излучения . . . . .	5
2.4	Решение RTE . . . . .	6
2.5	Пространственная симметрия . . . . .	7
<b>3</b>	<b>Машинное обучение</b>	<b>8</b>
3.1	Линейная регрессия или МНК . . . . .	8
3.2	Решающие деревья . . . . .	9
3.3	Ансамблевые методы. Случайный лес . . . . .	9
3.4	Бустинг . . . . .	10
<b>4</b>	<b>Практическая часть</b>	<b>10</b>

# 1 Введение

В современном мире методы машинного обучения играют все более значимую роль в различных научных исследованиях, включая область физики элементарных частиц. Одним из интересных направлений применения машинного обучения является моделирование распространения черенковских фотонов при детектировании нейтрино в эксперименте Baikal-GVD. Впервые идея создания нейтринного телескопа была предложена советским физиком-теоретиком М.А. Марковым в 1960-м году. Суть заключалась в регистрации неуловимых частиц в прозрачных природных средах с помощью детекторов света, расположенных на некотором расстоянии друг от друга. В 1980-м году при Институте ядерных исследований Российской академии наук (ИЯИ РАН) была основана Лаборатория нейтринной астрофизики высоких энергий, главной задачей которой стало создание телескопа для регистрации нейтрино в водах озера Байкал.

Эксперимент Baikal-GVD (Baikal Gigaton Volume Detector) представляет собой крупнейший в мире подводный нейтринный телескоп. Уникальная научная установка расположена в 3,6 километрах от берега на глубине около 1300 метров. Она состоит из десяти кластеров, каждый из которых содержит восемь вертикальных гирлянд (стрингов), прикрепленных ко дну. На каждой гирлянде закреплено по 36 оптических модулей - фотоэлектронных умножителей. Оптическая система этой установки предназначена для регистрации черенковского излучения мюонов и каскадов заряженных частиц высоких энергий, возникающих в результате нейтринных взаимодействий. Телескоп предназначен для изучения космических нейтрино через их взаимодействие с ледяной средой озера. Распространение черенковских фотонов при этом играет ключевую роль в регистрации и анализе сигналов.

Цель данной курсовой работы заключается в исследовании и применении методов машинного обучения для улучшения моделирования распространения черенковских фотонов в эксперименте Baikal-GVD. Акцент будет сделан на разработке алгоритмов, способных эффективно предсказывать потоки фотонов в среде, по характеристикам близкой к воде Байкала.

## 2 Теоретическая часть

### 2.1 Кинетические уравнения

Кинетические уравнения служат для описания макроскопических свойств вещества на основе молекулярно-кинетических представлений. С помощью кинетических уравнений можно рассчитать поведение газов, плаз-

мы, жидкостей при не слишком больших отклонениях от состояния термодинамического равновесия. В частности, кинетические уравнения используются для вычисления коэффициентов вязкости, теплопроводности, диффузии и т.д. Эта задача особенно актуальна сейчас, когда часто бывает нужно знать свойства вещества в недоступной для лабораторного эксперимента области изменения параметров.

Для описания распространения фотонов в среде используется уравнение переноса излучения, получаемое из уравнения Больцмана, которое, в свою очередь, является классическим примером кинетического уравнения. Далее представлена общая идея получения уравнения Больцмана для идеальных газов.

## 2.2 Уравнение Больцмана

Традиционным объектом рассмотрения кинетической теории являются идеальные газы. Для для их описания существует 2 подхода - макроскопический и микроскопический. На микроскопическом уровне наиболее полное описание одноатомного газа состоит в задании функции распределения всех атомов газа

$$f_N(x^1, v^1, x^2, v^2, \dots, x^N, v^N). \quad (1)$$

Эта функция с точностью до нормировки задает вероятность такого состояния газа, в котором атом 1 находится в точке  $x^1$  и имеет скорость  $v^1$  и т. д. Столь полное описание состояния газа бессмысленно ввиду очень большого числа атомов газа, и поэтому употребляются модели, использующие менее полную информацию. Одна из таких моделей основана на одночастичной функции распределения

$$f(x, v, t) = \int f_N(x, v, x^2, v^2, \dots, x^N, v^N t) dx^2 dv^2 \dots dx^N dv^N. \quad (2)$$

физический смысл функции  $f$  - это число частиц, имеющих скорость  $v$  и находящихся в окрестности точки  $x$ . Одно из центральных положений кинетической теории - утверждение о том, что одночастичная функция распределения удовлетворяет уравнению, из которого можно однозначно определить функцию  $f$  для всех значений времени  $t > 0$  исходя из ее значения в момент времени  $t = 0$ . Уравнение, из которого определяется функция  $f$ , называется кинетическим. В случае газов - это уравнение Больцмана.

Рассмотрим нестрогий вывод уравнения Больцмана.

Сначала для простоты мы предположим, что функция распределения  $f(x, v, t)$  не зависит от  $x$ :  $f(x, v, t) = F(v, t)$ . Будем считать, что газ состоит из одинаковых шариков диаметра  $\chi$ .

Пусть  $\mathbf{AB}$  - вектор, соединяющий центры двух атомов А и В в момент их столкновения. Очевидно, что  $\mathbf{AB} = \chi \mathbf{w}$ ,  $\mathbf{w}$  - вектор единичной длины. Пусть  $\mathbf{v}_A$  и  $\mathbf{v}_B$  - скорости атомов А и В до столкновения.

Найдём  $\mathbf{v}'_A$  и  $\mathbf{v}'_B$  после столкновения. Для этого запишем законы сохранения энергии и импульса.

$$\begin{aligned} \frac{1}{2}m(\mathbf{v}_A)^2 + \frac{1}{2}m(\mathbf{v}_B)^2 &= \frac{1}{2}m(\mathbf{v}'_A)^2 + \frac{1}{2}m(\mathbf{v}'_B)^2, \\ m\mathbf{v}_A + m\mathbf{v}_B &= m\mathbf{v}'_A + m\mathbf{v}'_B \end{aligned} \quad (3)$$

Обозначим приращение скорости через  $\mathbf{q}$ :

$$\begin{aligned} \mathbf{v}'_A &= \mathbf{v}_A + \mathbf{q}, \\ \mathbf{v}'_B &= \mathbf{v}_B - \mathbf{q} \end{aligned}$$

Имея  $\mathbf{n}$  - вектор единичной длины, подставляя в первое уравнение найдём

$$\mathbf{q} = \mathbf{n} \langle \mathbf{n}, (\mathbf{v}_B - \mathbf{v}_A) \rangle \quad (4)$$

При упругом ударе двух одинаковых шаров проекция их относительной скорости на линию центров меняет знак, а проекция на плоскость, перпендикулярную линии центров, сохраняется. Поэтому  $\mathbf{n} = \mathbf{w}$  и

$$\begin{aligned} \mathbf{v}'_A &= \mathbf{v}_A + \mathbf{w} \langle \mathbf{w}, (\mathbf{v}_B - \mathbf{v}_A) \rangle \\ \mathbf{v}'_B &= \mathbf{v}_B + \mathbf{w} \langle \mathbf{w}, (\mathbf{v}_B - \mathbf{v}_A) \rangle \end{aligned} \quad (5)$$

Подсчитаем число  $dN$  пар атомов, при столкновении которых вектор  $\mathbf{w}$  принадлежит элементу  $dS$  поверхности единичной сферы. Предположим, что вероятность атому А иметь скорость  $\mathbf{v}_A$  не зависит от вероятности атому В иметь скорость  $\mathbf{v}_B$ . Тогда

$$dN^- = \frac{1}{4\chi^2} F(\mathbf{v}_A, t) F(\mathbf{v}_B, t) \langle (\mathbf{v}_A - \mathbf{v}_B), \mathbf{w} \rangle dw dv_A dv_B dt \quad (6)$$

На величину  $dN^-$  уменьшится число пар атомов, скорости которых принадлежат малому параллелепипеду  $M$  в пространстве скоростей.

$$M = \{\mathbf{v}, \tilde{\mathbf{v}}; |\mathbf{v}_i - (\mathbf{v}_A)_i| \leq (dv_A)_i, |(\mathbf{v}_B)_i - \mathbf{v}_i| \leq (dv_B)_i, i = 1, 2, 3\}$$

Преобразование

$$\{\mathbf{v}_A, \mathbf{v}_B\} \rightarrow \{\mathbf{v}'_A, \mathbf{v}'_B\}$$

инволютивно, т.е. его повторное применение есть тождественное преобразование, и якобиан преобразования равен единице, поэтому число пар

атомов, скорости которых окажутся в множестве  $M$  после столкновений, будет равно

$$dN^+ = \frac{\chi^2}{4} F(\mathbf{v}'_A, t) F(\mathbf{v}'_B, t) \langle (\mathbf{v}_A - \mathbf{v}_B), \mathbf{w} \rangle dw dv_A dv_B dt. \quad (7)$$

Полное приращение числа пар атомов в множестве  $M$  за время  $dt$  в результате рассматриваемых столкновений будет равно

$$dN = \frac{1}{4\chi^2} [F(\mathbf{v}'_A, t) F(\mathbf{v}'_B, t) - F(\mathbf{v}_A, t) F(\mathbf{v}_B, t)] \langle (\mathbf{v}_A - \mathbf{v}_B), \mathbf{w} \rangle \times \quad (8) \\ \times dw dv_A dv_B dt.$$

Интегрируя по всем скоростям атомов  $B$ , направлениям вектора  $\mathbf{w}$  и вспоминая определение функции распределения, мы получим, что

$$\frac{\partial F}{\partial t} = \frac{1}{2\chi^2} \int [F(\mathbf{v}'_A, t) F(\mathbf{v}'_B, t) - F(\mathbf{v}_A, t) F(\mathbf{v}_B, t)] \langle (\mathbf{v}_A - \mathbf{v}_B), \mathbf{w} \rangle \times \quad (9) \\ \times dw dv_A dv_B dt$$

Это уравнение и есть уравнение Больцмана, записанное для рассматриваемого частного случая.

### 2.3 Уравнение переноса излучения

RTE (Radiative Transfer Equation) - это дифференциальное уравнение, описывающее излучение  $L(\mathbf{r}, \hat{\mathbf{s}}, t)$ .  $L$  можно трактовать как плотность потока энергии в пространственно-временной точке  $(\mathbf{r}, t)$  вдоль направления, заданного единичным вектором  $\hat{\mathbf{s}}$ . Уравнение учитывает потери энергии луча света из-за рассеяния и поглощения фотонов:

$$\frac{1}{c} \frac{\partial L(\mathbf{r}, t, \hat{\mathbf{s}})}{\partial t} + \hat{\mathbf{s}} \nabla L(\mathbf{r}, t, \hat{\mathbf{s}}) + \mu_t L(\mathbf{r}, t, \hat{\mathbf{s}}) = \quad (10) \\ = S(\mathbf{r}, t, \hat{\mathbf{s}}) + \mu_s \int_{S^2} d\hat{\mathbf{s}}_1 f(\hat{\mathbf{s}}', \hat{\mathbf{s}}) L(\mathbf{r}, t, \hat{\mathbf{s}}'),$$

где  $S(\mathbf{r}, t, \hat{\mathbf{s}})$  - функция, описывающая источник света,  $c$  - скорость света в среде,  $\mu_t = \mu_a + \mu_s$  - суммарный коэффициент затухания, обусловленный коэффициентами поглощения и рассеяния.

Интегрально-дифференциальное уравнение может быть преобразовано в интегральное уравнение:

$$L(\mathbf{r}, t, \hat{\mathbf{s}}) = L_0(\mathbf{r}, t, \hat{\mathbf{s}}) + \delta L(\mathbf{r}, t, \hat{\mathbf{s}}), \quad (11)$$

где

$$L_0 = c \int_0^t dt' e^{-c\mu_t(t-t')} S(\mathbf{r} - c\hat{\mathbf{s}}(t-t'), t', \hat{\mathbf{s}}),$$

$$\delta L = c\mu_s \int_{S^2} d\widehat{\mathbf{s}}_1 f(\widehat{\mathbf{s}}, \widehat{\mathbf{s}}') \int_0^t dt' e^{-c\mu_t(t-t')} L(\mathbf{r} - c\widehat{\mathbf{s}}(t-t'), t', \widehat{\mathbf{s}}). \quad (12)$$

Решение этого уравнения может быть найдено в виде

$$L^{(n)} = L_0 + \delta L^{(n-1)} \quad (13)$$

Запишем значения поправок для первых порядков.

$$\delta L^{(-1)} \equiv 0, L^{(0)} = L_0 \quad (14)$$

$$\delta L^{(1)} = c\mu_s \int_{S^2} d\widehat{\mathbf{s}}_1 f(\widehat{\mathbf{s}}, \widehat{\mathbf{s}}_1) \int_0^t dt_1 e^{-c\mu_t(t-t_1)} L_0(\delta\mathbf{r}, t_1, \widehat{\mathbf{s}}_1), \quad (15)$$

$$\begin{aligned} \delta L^{(2)} = & (c\mu_s)^2 \int_0^t dt_2 \int_0^{t_2} dt_1 e^{-c\mu_t(t-t_1)} \int_{S^2} d\widehat{\mathbf{s}}_1 f(\widehat{\mathbf{s}}, \widehat{\mathbf{s}}_2) \times \\ & \times \int_{S^2} d\widehat{\mathbf{s}}_1 f(\widehat{\mathbf{s}}, \widehat{\mathbf{s}}_1) L_0(\delta\mathbf{r}, t_1, \widehat{\mathbf{s}}_1), \end{aligned} \quad (16)$$

$$\begin{aligned} \delta L^{(n)} = & (c\mu_s)^n e^{-c\mu_t t} \left( \prod_{i=1}^n \int_0^{t_{i+1}} dt_i \right) e^{-c\mu_t t_1} \times \\ & \times \left( \prod_{j=1}^n \int_{S^2} d\widehat{\mathbf{s}}_{j+1} f(\widehat{\mathbf{s}}_j, \widehat{\mathbf{s}}_{j+1}) \right) L_0(\delta\mathbf{r}, t_1, \widehat{\mathbf{s}}_1). \end{aligned} \quad (17)$$

где  $\delta\mathbf{r} = \mathbf{r} - c\widehat{\mathbf{s}}(t-t_n) - c\widehat{\mathbf{s}}_n(t_n-t_{n-1}) \dots - c\widehat{\mathbf{s}}_2(t_2-t_1) - c\widehat{\mathbf{s}}_1 t_1$ .

## 2.4 Решение RTE

Для источника вида

$$S(\mathbf{r}, t, \widehat{\mathbf{s}}) = \delta^3(\mathbf{r})\delta(t)\delta^2(\widehat{\mathbf{s}} - \widehat{\mathbf{s}}_0) \quad (18)$$

можно показать, что полное решение будет следующим:

$$L = e^{-\mu_s c t} c \delta^3(\mathbf{r} - c\widehat{\mathbf{s}}_0 t) \delta^2(\widehat{\mathbf{s}} - \widehat{\mathbf{s}}_0) + e^{-\mu_s c t} \sum_{n=1}^{\infty} P_n(\mu_s c t) \delta L_s^{(n)} \quad (19)$$

где  $P_n = P_n(\mathbf{r}, t, \widehat{\mathbf{s}})$  - множитель, отвечающий за вероятность встретить фотон в точке  $\mathbf{r}$  с направлением  $\widehat{\mathbf{s}}$  в момент времени  $t$ . Этот множитель соответствует распределению Пуассона, которое моделирует появление  $n$  случайных некоррелированных рассеяний в течение интервала времени  $t$  в среднем с вероятностью  $c\mu_s$ , а  $L_n$  вычисляется по следующей формуле:

$$\begin{aligned} \delta L_s^{(n)} = & n! \prod_{i=1}^n \left( \int_0^1 d\xi_i (1 - \xi_i)^{i-1} \right) \prod_{j=1}^{n-1} \left( \int_{S^2} d\widehat{\mathbf{s}}_j f(\widehat{\mathbf{s}}_{j+1} \cdot \widehat{\mathbf{s}}_j) \right) \times \\ & \times f(\widehat{\mathbf{s}}_1 \cdot \widehat{\mathbf{s}}_0) \cdot c \delta^3(\mathbf{r} - c t \mathbf{s}^n) \end{aligned} \quad (20)$$

где

$$\mathbf{s}^n = \sum_{k=0}^n \widehat{\mathbf{s}}_{n-k} \xi_{n-k} \left( \prod_{i=1}^k (1 - \xi_{n-i+1}) \right), \quad (21)$$

$$\xi_k = 1 - \frac{t_k}{t_{k+1}}. \quad (22)$$

Первый член полученной формулы соответствует движению без рассеяния с затуханием скорости, а второй суммирует вклад рассеяний. В общем случае вычисление интегралов с  $n > 2$  выполняется численно.

## 2.5 Пространственная симметрия

Введём понятие средней координаты:

$$\mu'_s = \mu_s(1 - g), \langle \mathbf{r} \rangle = \frac{1}{\mu'_s} (1 - e^{-\mu'_s ct}) \widehat{\mathbf{s}}_0, \quad (23)$$

где  $g$  — параметр анизотропии среды.

Рассматривая приближение переноса, можно сделать следующие замечания:

Средняя координата никогда не превышает одну эффективную длину рассеяния, следовательно, на длинах, больших одной эффективной длины рассеяния, вероятность затеряться фотону стремится к 1.

Кроме того, средняя координата больше для  $g$  более близких к 1: если фотон рассеивается преимущественно вперед, то вероятность того, что он пройдет большее расстояние в этом направлении будет также больше.

Также заметим, что при малых временах средняя координата пропорциональна  $c \cdot t$ , что ожидаемо, когда мы работаем в приближении переноса. При больших временах в силу вступает диффузионное приближение и средняя координата перестает меняться.

Рассмотрим выражение для  $\delta L$ :

$$\begin{aligned} \delta L^{(n)} &= (c\mu_s)^n e^{-c\mu_s t} \left( \prod_{i=1}^n \int_0^{t_{i+1}} dt_i \right) \left( \prod_{j=1}^n \int_{S^2} d\widehat{\mathbf{s}}_j f(\widehat{\mathbf{s}}_{j+1}, \widehat{\mathbf{s}}_j) \right) \times \\ &\times f(\widehat{\mathbf{s}}_1 \cdot \widehat{\mathbf{s}}_0) \cdot c\delta^3(\mathbf{r} - ct\mathbf{s}^n), \end{aligned} \quad (24)$$

$$\begin{aligned} \delta G^{(n)} &= (c\mu_s)^n \left( \prod_{i=1}^n \int_0^{t_{i+1}} dt_i \right) \left( \prod_{j=1}^n \int_{S^2} d\widehat{\mathbf{s}}_j f(\widehat{\mathbf{s}}_{j+1}, \widehat{\mathbf{s}}_j) \right) \times \\ &\times f(\widehat{\mathbf{s}}_1 \cdot \widehat{\mathbf{s}}_0) \cdot c\delta^3(\mathbf{r} - ct\mathbf{s}^n). \end{aligned} \quad (25)$$

$$\delta G^{(n)}(\alpha r \mathbf{n}, l, \widehat{\mathbf{s}}), \delta^3(\alpha r \mathbf{n} - ct\widehat{\mathbf{s}}^{(n)}) = \frac{1}{\alpha^3} (r \mathbf{n} - \frac{ct}{\alpha} \widehat{\mathbf{s}}^{(n)}), \quad (26)$$

откуда

$$\delta G^{(n)}(\alpha r \mathbf{n}, l, \hat{\mathbf{s}}) = \alpha^{n-3} \delta G^{(n)}(r \mathbf{n}, \frac{l}{\alpha}, \hat{\mathbf{s}}). \quad (27)$$

Поэтому для  $\delta L^{(n)}$  справедливо выражение:

$$\delta L^{(n)}(\alpha r \mathbf{n}, l, \hat{\mathbf{s}}) = e^{-\mu_t l \frac{\alpha-1}{\alpha}} \alpha (n-3) \delta L^{(n)}(r \mathbf{n}, \frac{l}{\alpha}, \hat{\mathbf{s}}). \quad (28)$$

Полученное соотношение даёт возможность получить поправку порядка  $n$  в момент времени  $t$  в точке, находящейся на расстоянии  $\alpha r$  от источника, при условии, что известна поправка этого порядка для точки, находящейся на расстоянии  $r$  в момент времени  $\frac{t}{\alpha}$ .

## 3 Машинное обучение

Как уже говорилось ранее, в работе применяются методы машинного обучения для моделирования потока фотонов в некоторой точке пространства на заданном временном промежутке. В этой части представлена теория, связанная с принципами работы используемых моделей.

### 3.1 Линейная регрессия или МНК

Линейная регрессия является одним из простейших и наиболее распространенных методов машинного обучения, используемых для предсказания значений зависимой переменной на основе одной или нескольких независимых переменных. В линейной регрессии предполагается, что зависимость между переменными линейна, то есть можно описать уравнением прямой линии.

Основная цель линейной регрессии - найти такие коэффициенты (веса), которые минимизируют сумму квадратов разностей между реальными значениями зависимой переменной и предсказанными моделью. Для этого часто используется метод наименьших квадратов.

Линейная регрессия может быть применена как для простых случаев с одной независимой переменной (простая линейная регрессия), так и для более сложных ситуаций с несколькими независимыми переменными (множественная линейная регрессия).

Преимущества линейной регрессии включают простоту интерпретации результатов, относительную быстроту обучения модели и хорошую обобщающую способность. Однако, недостатком может быть то, что модель предполагает линейную зависимость между переменными, что может быть недостаточно для точного предсказания в сложных сценариях.

## 3.2 Решающие деревья

Решающие деревья в машинном обучении представляют собой модель, которая используется для прогнозирования целевой переменной на основе входных признаков. Суть работы решающих деревьев заключается в построении древовидной структуры, где каждый узел представляет собой условие, которое разделяет данные на две или более подгруппы. Эти условия выбираются таким образом, чтобы максимизировать однородность подгруппы по целевой переменной.

Процесс построения решающего дерева начинается с корневого узла, который содержит все обучающие данные. Затем алгоритм делит данные на две подгруппы в соответствии с выбранным условием. Этот процесс повторяется для каждой подгруппы, пока не будет достигнут критерий остановки (например, максимальная глубина дерева, минимальное количество объектов в узле и т.д.).

После построения дерева можно использовать его для прогнозирования новых данных, просто проходя по дереву от корня к листьям и применяя условия для определения принадлежности нового объекта к определенному классу или значению целевой переменной.

Решающие деревья обладают простотой интерпретацией, способностью обрабатывать как категориальные, так и числовые признаки, а также автоматическим отбором признаков. Однако они могут быть склонны к переобучению на обучающих данных и не всегда являются самыми точными моделями.

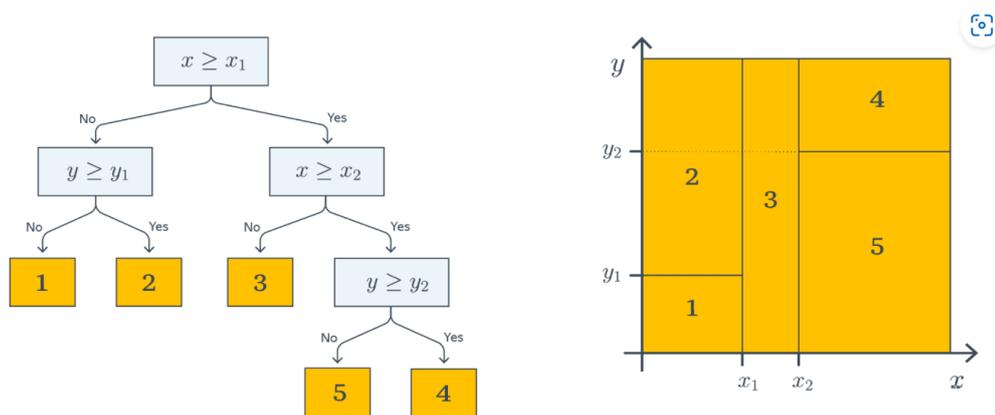


Рис. 1: Иллюстрация работы алгоритма решающего дерева

## 3.3 Ансамблевые методы. Случайный лес

Ансамбли в машинном обучении представляют собой методы, которые объединяют несколько моделей, называемых базовыми, для улучшения

качества прогнозов. Они основаны на идее, что объединение нескольких слабых моделей может привести к созданию более сильной и устойчивой.

Один из популярных ансамблевых методов - случайный лес. Суть работы случайного леса заключается в построении нескольких решающих деревьев и объединении их результатов для получения окончательного прогноза. Каждое дерево строится на случайной подвыборке данных (bootstrap sample) и случайном подмножестве признаков (feature bagging), что делает модели разнообразными и уменьшает вероятность переобучения.

При прогнозировании новых данных каждое дерево в случайном лесу выдает свой собственный прогноз, и окончательное решение принимается путем голосования или усреднения результатов всех деревьев. Таким образом, случайный лес уменьшает дисперсию прогнозов и повышает обобщающую способность модели.

Преимущества случайного леса включают в себя высокую точность прогнозов, способность обрабатывать большое количество признаков и автоматический отбор наиболее важных признаков. Кроме того, случайный лес устойчив к переобучению и шуму в данных.

Однако следует отметить, что случайный лес может быть затратным по вычислительным ресурсам из-за большого количества деревьев, а также менее интерпретируемым по сравнению с отдельными решающими деревьями.

### 3.4 Бустинг

Бустинг (boosting) — это ансамблевый метод, в котором строится множество базовых алгоритмов из одного семейства, объединяющихся затем в более сильную модель. Отличие состоит в том, что в случайном лесе базовые алгоритмы учатся независимо и параллельно, а в бустинге — последовательно. Каждый следующий базовый алгоритм в бустинге обучается так, чтобы уменьшить общую ошибку всех своих предшественников. Как следствие, итоговая композиция будет иметь меньшее смещение, чем каждый отдельный базовый алгоритм.

## 4 Практическая часть

При вычислении интегралов с помощью VEGAS Monte-Carlo была получена зависимость  $\Delta N_{hit}$  от  $t$ , напоминающую логарифмическую функцию. Поэтому появилась идея сделать аппроксимацию на основе выбранных сеточных значений, использование которой могло бы значительно сократить время работы программы.

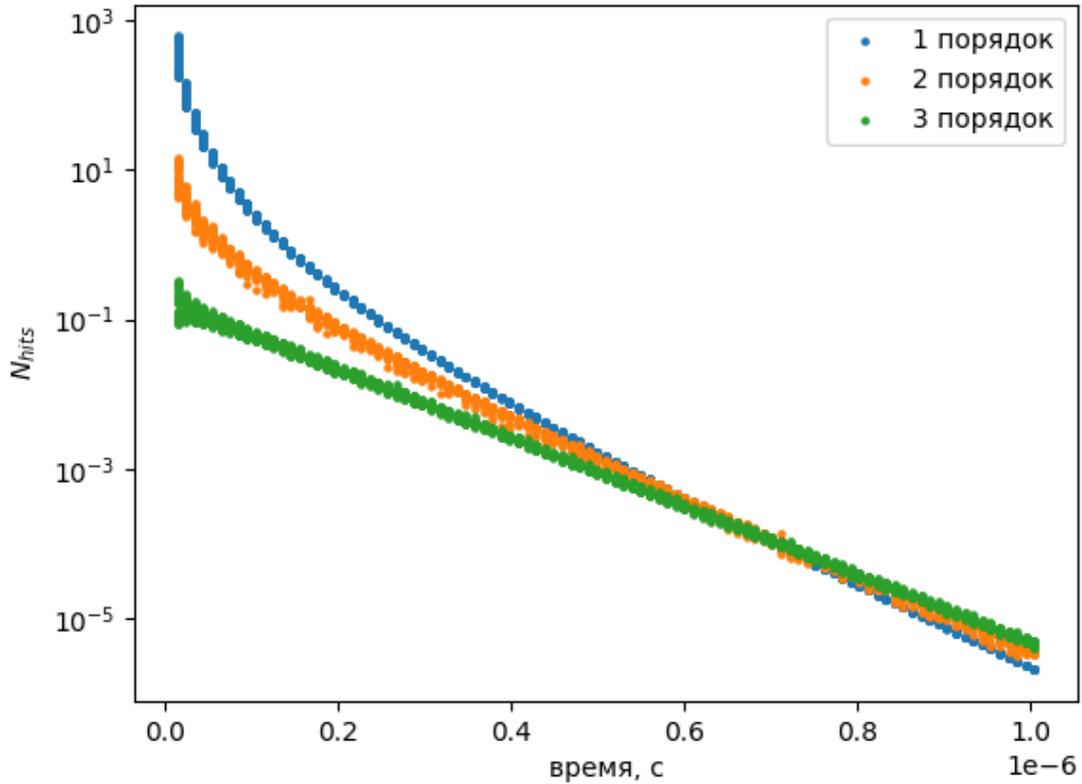


Рис. 2: Ожидаемое число детектируемых фотонов  $\Delta N_{hit}$  (первые 3 порядка), полученных при решении уравнения переноса излучения

Для решения задачи аппроксимации в данной работе используются и сравниваются следующие методы машинного обучения:

- линейная регрессия
- решающее дерево
- случайный лес
- градиентный бустинг

Для оценки качества модели используется метрика  $R^2$ . Эта метрика является одной из ключевых для оценки качества моделей в машинном обучении.  $r^2$  показывает, насколько хорошо модель соответствует данным и объясняет вариацию целевой переменной.

Значение  $r^2$  может варьироваться от 0 до 1 и интерпретируется следующим образом:

- $r^2 = 0$ : модель не объясняет никакой вариации данных, предсказания равны среднему значению целевой переменной.
- $r^2 = 1$ : модель идеально соответствует данным, объясняя всю вариацию целевой переменной.

Итак, для обучения моделей был собран датасет на следующей сетке:

$$R = 1; \phi = 0; \theta = 0.1, 0.2, \dots, \pi;$$

$$t = \left( \frac{R}{c}; \frac{R}{c} + 100 \cdot 10^{-9} \right) \cup \left( \frac{R}{c} + 100 \cdot 10^{-9}; \frac{R}{c} + 100 \cdot 10^{-6} \right)$$

На каждом временном интервале было выбрано 100 точек.

Используется единственное сеточное значение  $R = 1$  т. к. в задаче имеет место пространственная симметрия, о которой уже шла речь ранее. Для получения значений с произвольным  $R$  предполагается использовать формулу (28).

Плоскость рассеяния фотонов была выбрана следующим образом: угол  $\phi$  положен равным 0, а для  $\theta$  выбран интервал от 0 до  $\pi$ . Так как задача распространения света от точечного источника симметрична относительно центра этого источника, то очевидно, что схожие результаты были бы получены при любом другом значении  $\phi$ . Также важно заметить, что для упрощения задачи рассеяния 2, 3 и 4 порядков рассматриваются в одной и той же плоскости, несмотря на то, что рассеяния с  $n > 2$  обычно отклоняются от плоскости рассеяния.

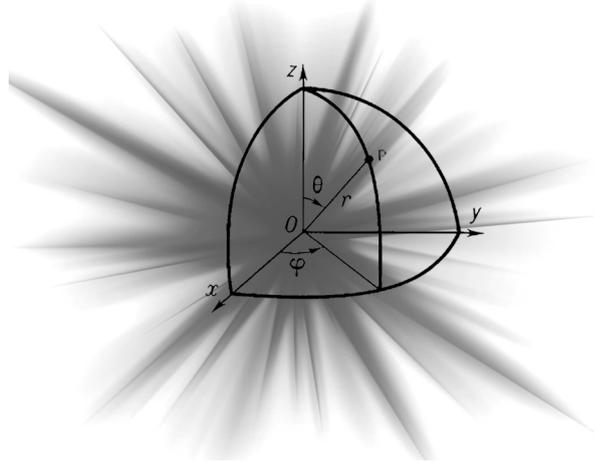


Рис. 3: Распространения света от точечного источника

Работа началась со сбора данных. С помощью интегрирования были получены значения  $\Delta N_{hit}$  для первых 4 порядков в указанных сеточных значениях. Для обучения данные были предобработаны:

- удалены значения  $\Delta N_{hit} < 10^{-100}$ ;
- получены значения  $\ln(\Delta N_{hit})$ , подходящие для обучения на них линейных моделей;
- значения времени из [с] переведены в [нс];
- параметр  $g$  заменён на  $(1-g)$ .

С помощью таких библиотек, как `sklearn.ensemble`, `sklearn.model_selection`, и `sklearn.metrics` данные были разделены на тренировочную и тестовую часть. Затем на тренировочной части данных были обучены несколько моделей ML, после чего их работа проверена на тестовой. Оказалось, что все модели хорошо справляются с поставленной задачей на имеющихся

данных, вследствие чего их можно использовать без потери точности. В таблице 1 приведено сравнение значений  $R^2$  для различных моделей.

метод	$R^2$
linear regression	0.9822856702171563
decision tree	0.9987289116771565
random forest	0.9987557341290492
gradient boosting	0.9997281712582555

Таблица 1: Сравнение  $R^2$  для различных моделей ML

Также была реализована функция, выполняющее "растяжение" в соответствии с формулой (28). Для получения значений  $\Delta N_{hit}$ , соответствующих времени  $\frac{t}{\alpha}$  была использована интерполяция, реализованная с помощью библиотеки `scipy.interpolate`.

Проверим работу полученной программы на примере конкретной точки с

$$R = 5; \phi = 0; \theta = \frac{\pi}{2}.$$

и том же временном промежутке.

На графике представлено сравнение результатов, полученных с помощью решения RTE в данной точке, а также полученных с помощью машинного обучения и в последствие растянутых на нужный радиус.

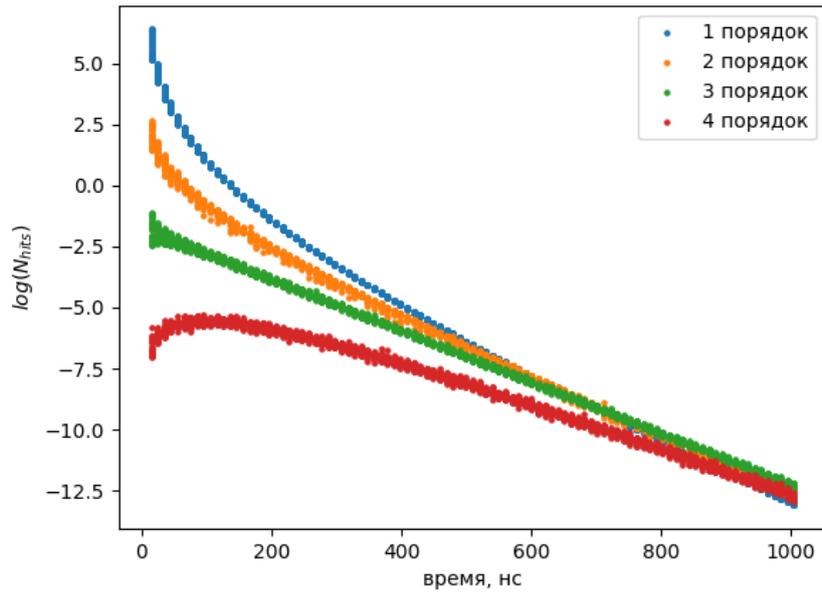


Рис. 4: Данные после обработки

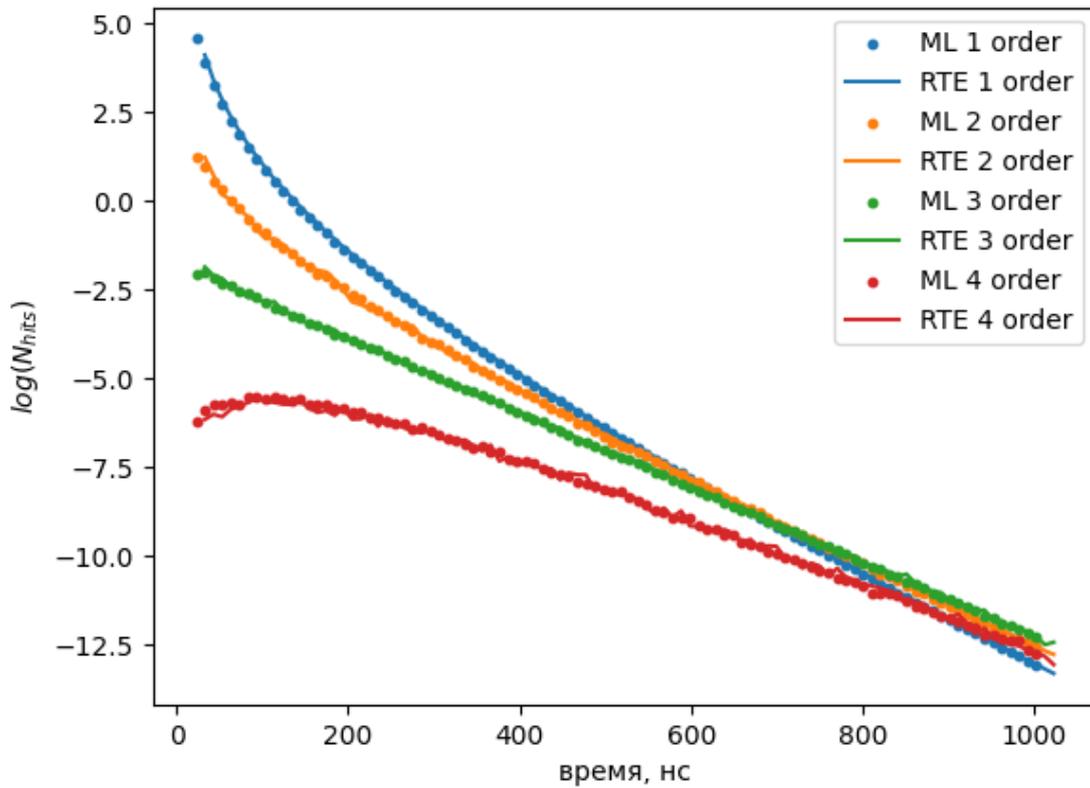


Рис. 5: Сравнение результатов, полученных с помощью решения RTE и с помощью машинного обучения (ML), в последствие растянутых на  $R = 5$ .

## Заключение

Проведён анализ основ теории кинетических уравнений и, в частности, теории, связанной с распространением фотонов в среде, способной к рассеиванию и поглощению, составлен соответствующий литературный обзор.

Также изучены такие методы машинного обучения, как линейная регрессия, решающее дерево, случайный лес и градиентный бустинг; сведения о принципах их работы изложены в литературном обзоре. Практическое применение даёт возможность достаточно точно предсказывать поток фотонов в произвольной точке пространства, о чём свидетельствует метрика оценки качества  $R^2$ , которая для каждой модели оказалась близка к 1.

Ещё одним свидетельством правомерности метода является корректная работа разработанной программы для произвольной точки.

## Список литературы

- [1] Арсеньев А.А. Кинетические уравнения. — М.: Знание, 1985, — (Новое в жизни, науке, технике. Сер. "Математика, кибернетика"; №1), — 64 с.
- [2] Смелов В.В., Лекции по теории переноса нейтронов. — М.: Атомиздат, 1978, — (Изд. 2-е, перераб. и дополн.),— 360 с.
- [3] Vladimir Allakhverdian, Dmitry V. Naumov. Infinite Series Solution of the Time-Dependent Radiative Transfer Equation in Anisotropically Scattering Media. — JINR, Dubna, Russia 141980, 2024, —(Journal of Quantitative Spectroscopy and Radiative Transfer January 29)