

НЕЙРОННЫЕ ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ: КОГДА ГЛУБОКОЕ ОБУЧЕНИЕ ВСТРЕЧАЕТСЯ С МАТЕМАТИЧЕСКИМ МОДЕЛИРОВАНИЕМ

Цзи Пэнбо, студент

Московский государственный технический университет им. Н.Э. Баумана
(Россия, г. Москва)

DOI:10.24412/2500-1000-2025-3-1-344-349

Аннотация. В данной статье рассматриваются нейронные дифференциальные уравнения, когда глубокое обучение встречается с математическим моделированием. Актуальность темы обусловлена тем, что современные технологии машинного обучения, в частности глубокое обучение, находятся на пике своего развития и находят применение в самых различных областях науки и техники. Однако многие особенности их использования и обоснования остаются недостаточно изученными, что затрудняет оптимизацию и расширение области их применения. Целью данного исследования является разработка математической модели нейронных дифференциальных уравнений для анализа и предсказания динамики нейронных сетей во времени. Для достижения цели необходимо решить задачи, направленные на изучение существующих подходов к математическому моделированию нейронных сетей и выявление их связи с дифференциальными уравнениями, а также на применение численных методов для решения полученной модели с исследованием поведения нейронной сети при различных входных сигналах. В качестве методов используются математическое моделирование, численные методы, сравнительный анализ. Теоретическая значимость заключается в расширении представлений о математическом моделировании нейронных сетей, а также в разработке новых математических моделей, учитывающих сложность динамических процессов, происходящих в сетях глубокого обучения. Практическая значимость заключается в разработке математической модели нейронных дифференциальных уравнений и описании динамики нейронных сетей во времени. Разработанные модели могут быть применены для улучшения качества предсказаний, оптимизации процессов обучения и работы нейронных сетей в реальных условиях.

Ключевые слова: нейронные сети, дифференциальные уравнения, глубокое обучение, математическое моделирование, алгоритмы машинного обучения, нейронные дифференциальные уравнения.

Дифференциальные уравнения служат основой для описания динамики многих физических, биологических, экономических и социальных систем. Они позволяют предсказать поведение систем на основе их начальных условий. В некоторых случаях традиционные методы, основанные на аналитических решениях, сталкиваются с трудностью из-за сложности описания процессов, особенно в условиях ограниченности данных. Современные нейронные сети открывают новые возможности для моделирования таких систем, позволяя эффективно сочетать машинное обучение и математическое моделирование. Одним из таких подходов является использование нейронных дифференциальных уравнений (НДУ), которые необходимы для работы с динамическими системами, где традиционные методы не всегда подходят [2, с. 75].

Дифференциальные уравнения – естественный язык физики и природы. В школе нам объясняли второй закон Ньютона о движении: $F=ma$. Ускорение a представляет собой вторую производную позиции x . Следовательно, $F=mx''$, что является дифференциальным уравнением второго порядка. Благодаря своей эффективности дифференциальные уравнения нашли широкое применение в биологии, экономике и социологии. Методы машинного обучения доказали свою способность успешно предсказывать поведение сложных систем.

НДУ представляют собой гибридный подход, где традиционные математические модели дифференциальных уравнений интегрируются с мощью нейронных сетей. Они открывают новые горизонты для задач, требующих описания динамики сложных систем. Глубо-

кое обучение встречается с математическим моделированием, когда необходимо аппроксимировать сложные функции или векторные поля, которые описывают поведение динамических систем.

В отличие от традиционных методов, которые часто требуют знания физики системы или сложных аналитических подходов для

где $z(t)$ – состояние системы;
 $f(t, z(t), \theta)$ – это векторное поле, описывающее динамику системы;
 t – время;
 θ – параметры, которые определяют свойства системы.

где $z_1(t) := \theta$ – угол отклонения маятника от вертикали;
 $z_2(t) := \dot{\theta}$ – угловая скорость маятника;
 g – гравитационная постоянная;
 L – длина маятника.

Это дифференциальное уравнение описывает изменения угла и угловой скорости маятника во времени. В традиционном подходе для решения таких уравнений использовались численные методы интегрирования. В более сложных системах, где поведение не может

где $NN(t, z(t), \theta_{NN})$ – нейронная сеть, обучаемая на данных для аппроксимации динамики системы.

В НДУ нейронная сеть играет роль функции, описывающей динамику системы.

Рисунок 1 представляет собой схематическое изображение маятника, который является

вычисления параметров, нейронные сети могут работать с данными напрямую. Это позволяет моделировать системы, для которых нет явных аналитических решений или для которых сложно собрать необходимые теоретические данные.

В общем случае дифференциальное уравнение может быть записано в таком виде:

$$z'(t) = f(t, z(t), \theta) \quad (1)$$

Дифференциальные уравнения моделируют скорость изменения переменных во времени. Для примера рассмотрим дифференциальное уравнение, описывающее движение простого маятника. Его динамика описывается системой уравнений:

$$z'(t) = [z_2(t) - g \sin(z_1(t))/L] \quad (2)$$

быть описано явными уравнениями, требуется использование таких методов машинного обучения, как нейронные сети для аппроксимации векторного поля [4, с. 183].

В отличие от традиционных методов, когда у нас есть только данные о входах и выходах системы, можно применить нейронные сети для аппроксимации векторного поля f . Нейронные сети, используемые для аппроксимации векторного поля, преобразуют дифференциальное уравнение в форму:

$$z'(t) = NN(t, z(t), \theta_{NN}) \quad (3)$$

примером системы, описываемой дифференциальным уравнением. Используется маятник для иллюстрации перехода от классического подхода к нейронным дифференциальным уравнениям.

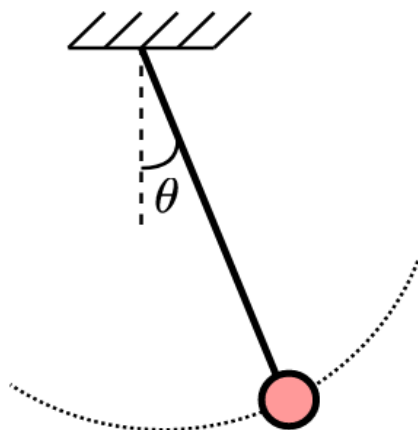


Рис. 1. Схематический рисунок маятника

Таким образом, эта простая модель помогает продемонстрировать основные принципы

работы с дифференциальными уравнениями и их применение в различных областях.

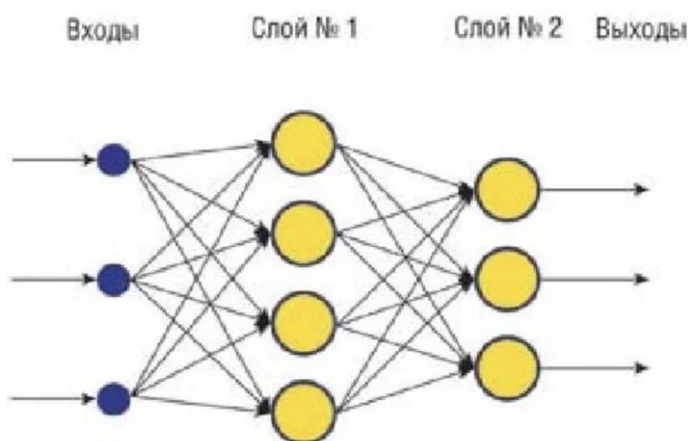


Рис. 2. Модель НДУ

Нейронная сеть получает начальные входные условия для решения дифференциального уравнения. Эти данные находятся с левой стороны модели и являются отправной точкой для вычисления дальнейших изменений во времени.

Нейронная сеть находится в центре и представлена как несколько слоев. Она принимает входные данные и использует функции активации для обработки информации. На каждом слое нейронная сеть генерирует аппроксимацию векторного поля, который описывает динамику системы. Сеть учится через обучение на данных, что позволяет ей выявлять скрытые закономерности и моделировать поведение системы.

На выходе нейронной сети формируется динамическое изменение состояния системы. Это представление находится с правой стороны модели и связано с динамикой системы во

времени, где каждый выходной вектор соответствует состоянию системы в определенный момент времени. Выходные данные используются для предсказания дальнейших изменений состояния системы на основе обученной модели.

На модели показано, как традиционные численные методы решения дифференциальных уравнений (метод Эйлера или Рунге-Кутты) используют шаги времени и аппроксимируют решения через дискретизацию, в то время как нейронная сеть пытается выучить непрерывное представление системы.

При решении практической задачи предполагаем, что у нас есть нейронная сеть с одним скрытым слоем, которая моделирует динамику выхода в зависимости от входных данных. Моделируем её с использованием нейронных дифференциальных уравнений, где актив-

ность нейронов в скрытом слое изменяется во времени.

Уравнение нейрона в скрытом слое можно описать следующим образом:

$$dh(t) / dt = -\alpha h(t) + w_1 x(t) + w_2 h(t), \quad (4)$$

где $h(t)$ – скрытая активация нейрона в момент времени t ,

α – коэффициент, определяющий скорость затухания нейрона,

w_1, w_2 – веса нейрона,

$x(t)$ – входной сигнал в момент времени t .

Вектор входных данных $x(t)$ и выходных данных $y(t)$ могут быть представлены как функции времени.

Нам необходимо найти решение для скрытых активаций нейрона $h(t)$, а затем использовать её для вычисления выхода нейронной сети $y(t)$.

Уравнение $dh(t)/dt = -\alpha h(t) + w_1 x(t) + w_2 h(t)$ – это линейное уравнение с постоянными коэффициентами. Для его решения мы можем использовать метод интегрирования через переменные разложения (или численные методы, такие как метод Эйлера или Рунге-Кутты).

Вход $x(t)$ – это заданная функция, например, $x(t)=\sin(t)$

Предположим, что $\alpha=0.5$, $w_1=0.8$, $w_2=0.3$, и начальное условие $h(0)=0$. Применяем метод Эйлера для численного решения уравнения на отрезке времени от $t=0$ до $t=T$.

Используем шаг по времени dt :

$$h(t+dt) = h(t) + (-\alpha h(t) + w_1 x(t) + w_2 h(t)) dt. \quad (5)$$

Выход сети $y(t)$ зависит от активации скрытого слоя через некоторую нелинейность (например, сигмоиду):

$$y(t) = 1 / (1 + e^{-h(t)}). \quad (6)$$

Решив систему, мы получаем зависимость выходных данных $y(t)$ от входных $x(t)$, что позволяет анализировать, как динамика нейронной сети зависит от входных сигналов и параметров модели.

Входной сигнал $x(t)=\sin(t)$, а параметры: $\alpha = 0,5$, $w_1 = 0,8$, $w_2 = 0,3$.

Используя метод Эйлера с шагом $dt = 0,1$, получаем решение для $h(t)$ и затем для $y(t)$.

Появляется возможность моделировать поведение нейронной сети во времени с помощью дифференциальных уравнений.

Использование нейронных сетей в качестве аппроксиматоров для векторных полей в дифференциальных уравнениях открывает новые горизонты для математического моделирования. Применение нейронных дифференциальных уравнений позволяет улучшить точность предсказаний и оптимизировать процессы обучения нейронных сетей в реальных условиях, где традиционные методы не могут быть применены или сталкиваются с проблемами из-за сложности или недостатка данных. Нейронные дифференциальные уравнения позволяют эффективно работать с системами, где традиционные аналитические

методы не могут быть применены [7, с. 64]. Преимущества этого подхода включают в себя способность работать с большими объемами данных, адаптацию к неравномерным временным рядам и уменьшение вычислительных затрат за счет меньшего количества параметров по сравнению с другими глубокими нейронными сетями. НДУ могут существенно улучшить точность и эффективность моделей, что имеет огромный потенциал для их использования в научных исследованиях, инженерии, биологии, экономике и других областях, где традиционные методы моделирования ограничены.

НДУ открывают множество новых возможностей в математическом моделировании, особенно когда речь идет о сложных системах, где традиционные методы моделирования ограничены. Однако их использование сопряжено с рядом проблем, которые требуют поиска эффективных решений. Одной из главных проблем является высокая вычислительная сложность обучения таких моделей. Нейронные сети требуют значительных вычислительных ресурсов. Процесс тренировки НДУ может занять много времени, особенно

когда данные имеют высокую временную разрешающую способность. Чтобы решить эту проблему, можно использовать такие методы оптимизации, как стохастический градиентный спуск с адаптивными алгоритмами обучения. Необходимо рассматривать использование распределенных вычислений или облачных сервисов для ускорения тренировочного процесса [5, с. 4]. Можно уменьшить сложность сети путем оптимизации ее архитектуры.

Другой значимой проблемой является трудность интерпретации параметров модели. Нейронные сети являются «черными ящиками», и понимание того, как различные параметры влияют на решение задачи, может быть затруднено. Иногда нужно объяснить поведение модели в физике или других дисциплинах, где интерпретируемость имеет большое значение. Одним из решений этой проблемы является использование методов объяснимого машинного обучения, таких как SHAP или LIME. Эти методы помогают интерпретировать поведение нейронных сетей и объяснить, как они принимают решения. Развитие теоретических подходов к нейронным сетям и их параметрам позволит повысить интерпретируемость таких моделей.

С проблемой неравномерной выборки данных сталкиваются НДУ при работе с временными рядами. В реальных приложениях данные часто бывают представлены с различными временными интервалами. Однако НДУ имеют явные преимущества в этом случае, поскольку они способны адаптировать временную сетку интегратора. Для повышения точности можно использовать адаптивные интеграторы или методы интерполяции для более точной аппроксимации динамики системы. В результате улучшается модель при работе с нерегулярными данными.

Проблемы в обобщении моделей на новые данные возникают, если тренировочный набор ограничен или не включает все воз-

можные вариации данных. Нейронные сети могут плохо обобщать на новые данные, что снижает точность прогнозов [6, с. 31]. Чтобы решить эту проблему, можно использовать методы регуляризации, которые позволяют избежать переобучения, методы аугментации данных, которые увеличивают объем тренировочного набора за счет создания дополнительных вариаций данных. Необходимо применять кросс-валидацию и тестирование на независимых данных для проверки способности модели адаптироваться к новым ситуациям.

Существует проблема недостатка данных для тренировки модели [3, с. 182]. Для преодоления этой проблемы можно использовать методы трансфера обучения, когда модель обучается на данных, полученных из другой, более близкой задачи. Необходимо применять синтетическую генерацию данных с помощью генеративных моделей. Техники активного обучения могут повысить эффективность использования малых данных и позволяют улучшать модель с каждым новым наблюдением.

Таким образом, нейронные дифференциальные уравнения представляют собой мощный инструмент для улучшения качества предсказаний и оптимизации процессов обучения нейронных сетей. Их использование в реальных задачах позволяет преодолеть ограничения традиционных методов и значительно улучшить результаты моделирования в различных областях. Несмотря на значительные преимущества нейронных дифференциальных уравнений в моделировании сложных систем, необходимо учитывать и решать различные проблемы. Применение методов оптимизации вычислений, улучшение интерпретируемости, работа с неравномерными данными и использование техник активного обучения и трансфера данных помогут преодолеть проблемы и повысить эффективность НДУ в реальных задачах.

Библиографический список

1. Арабов М.К. Применение нейронных сетей для решения дифференциальных уравнений и их анализ // Роль филиала Московского государственного университета имени М.В. Ломоносова в городе Душанбе в развитии науки и образования. – Душанбе, 2024. – С. 7-11.
2. Ахметов И.З. Сравнение классических методов численного решения дифференциальных уравнений с методом нейронных сетей // Современная наука: актуальные проблемы теории и практики. – 2024. – № 8. – С. 71-76.

3. Мамедалиев Р.Р. Интеграция дифференциальных уравнений и нейронных сетей для прогнозирования временных рядов и систем с обратной связью // Компьютерное проектирование в электронике. – Минск, 2024. – С. 180-184.
4. Музалевский А.В., Репин С.И. О решении дифференциальных уравнений с помощью нейронных сетей: целевые функционалы и верификация результатов // Математическое моделирование. – 2024. – Т. 36. № 6. – С. 179-200.
5. Петров Г.Г. Роль метода дифференциального уравнения для обучения нейронной сети // Уральский научный вестник. – 2023. – Т. 9. № 5. – С. 3-4.
6. Семькина Н.А., Тишина Е.В. Применение дифференциальных уравнений в моделировании нейронных сетей // Безопасность. Управление. Искусственный интеллект. – 2024. – Т. 3. № 3 (3). – С. 28-32.
7. Слияков П.А. Нейронные обыкновенные дифференциальные уравнения и возможность их аналитического решения // Аналитические методы анализа и дифференциальных уравнений. Материалы 11-го международного научного семинара. – Минск, 2024. – С. 63-64.

NEURAL DIFFERENTIAL EQUATIONS: WHEN DEEP LEARNING MEETS MATHEMATICAL MODELING

Ji Pengbo, *Student*

Bauman Moscow State Technical University
(Russia, Moscow)

Abstract. *This article discusses neural differential equations when deep learning meets mathematical modeling. The relevance of the topic is due to the fact that modern machine learning technologies, in particular deep learning, are at the peak of their development and find application in a wide variety of fields of science and technology. However, many features of their use and justification remain insufficiently studied, which makes it difficult to optimize and expand the scope of their application. The purpose of this study is to develop a mathematical model of neural differential equations for analyzing and predicting the dynamics of neural networks over time. To achieve this goal, it is necessary to solve problems aimed at studying existing approaches to mathematical modeling of neural networks and identifying their relationship to differential equations, as well as applying numerical methods to solve the resulting model by studying the behavior of a neural network with various input signals. Mathematical modeling, numerical methods, and comparative analysis are used as methods. The theoretical significance lies in expanding the understanding of mathematical modeling of neural networks, as well as in the development of new mathematical models that take into account the complexity of dynamic processes occurring in deep learning networks. The practical significance lies in the development of a mathematical model of neural differential equations and the description of the dynamics of neural networks over time. The developed models can be used to improve the quality of predictions, optimize learning processes and the operation of neural networks in real conditions.*

Keywords: *neural networks, differential equations, deep learning, mathematical modeling, machine learning algorithms, neural differential equations.*