Факультет физико-математических и естественных наук

Кафедра математического моделирования и искусственного интеллекта

Москва, 2024

Российский университет дружбы народов

Обучение нейронных сетей для аппроксимации решений краевых задач с применением неклассических вариационных формулировок

Студент:

Габриэль Тьерри

Группа: **НКНбд-01-20**

Номер билета: **1032204249**

Научный руководитель:

Шорохов Сергей Геннадьевич,

к.ф.-м.н., доцент

Введение

- **Актуальность работы**:

Проблемы современных методов: Современные методы решения сложных задач часто оказываются слишком медленными и сложными.

Преимущества нейронных сетей: Нейронные сети предлагают новый и более эффективный способ, позволяющий решать задачи быстрее и точнее.

- **Цель исследования**:

Изучение возможностей нейронных сетей в решении краевых задач.

- **Объект исследования**:

Нейронные сети для аппроксимации решений краевых задач.

- **Предмет исследования**:

Применение неклассических вариационных формулировок.

Задачи работы

Для достижения цели, в работе решаются следующие задачи:

- 1. Исследование создания вариационной формулировки уравнений колебаний конечной струны
- 2. Разработка архитектуры нейронной сети (PINN), которая включает в себя слои, подходящие для понимания и обработки динамики краевых задач, такие как слои долговременной кратковременной памяти (LSTM) и плотные слои.
- 3. Внедрение специального процесса обучения, который корректировал веса сети в зависимости от ее эффективности в прогнозировании точных решений и соблюдении физических законов, используя методы оптимизации на основе градиента.
- 4. Использование визуального метода для отображения прогнозов нейронной сети и сравнения их с точными решениями обеспечивает четкую визуальную оценку точности нейронной сети и качества ее аппроксимаций

Структура работы

- **Первый раздел:**
  Дается обзор дифференциальных уравнений, методов их решения и вариационных принципов.

- **Второй раздел:**
  Описывается использование нейронных сетей для аппроксимации решений дифференциальных уравнений. Представлена задача создания неклассической вариационной формулировки для одномерного волнового уравнения по методу В.М. Шалова. Вариационный функционал преобразован для удобства использования метода Монте-Карло.

- **Третий раздел:**
  Рассматривается архитектура нейронной сети для аппроксимации решений краевых задач для волнового уравнения. Приводятся детали реализации и обучения сети. Описаны результаты вычислительного эксперимента, подтверждающие высокую точность аппроксимации.

постановка задания

Рассмотрим построение вариационной формулировки для уравнений колебаний конечной струны

$$u_{tt} - u_{xx} = 0 \qquad (1)$$

в квадратной области $\Omega = [0, \pi] \times [0, \pi]$. Граница $\partial\Omega$ области Ω является объединением четырех отрезков $\gamma 1$, $\gamma 2$, $\gamma 3$, $\gamma 4$

$$\partial\Omega = \gamma 1 \cup \gamma 2 \cup \gamma 3 \cup \gamma 4,$$

Зададим следующие граничные условия на $\partial\Omega$

$$\begin{cases} u\,|_{\gamma_1} = 0,\ 0 \leqslant x \leqslant \pi,\ t = 0, \\ u_x\,|_{\gamma_2} = -\sin t,\ x = \pi,\ 0 \leqslant t \leqslant \pi, \\ u_t\,|_{\gamma_2} = 0,\ x = \pi,\ 0 \leqslant t \leqslant \pi, \\ u_x\,|_{\gamma_3} = 0,\ 0 \leqslant x \leqslant \pi,\ t = \pi, \\ u_t\,|_{\gamma_3} = -\sin x,\ 0 \leqslant x \leqslant \pi,\ t = \pi, \\ u_x\,|_{\gamma_4} = \sin t,\ x = 0,\ 0 \leqslant t \leqslant \pi \\ u_t\,|_{\gamma_4} = 0,\ x = 0,\ 0 \leqslant t \leqslant \pi. \end{cases} \qquad (2)$$

Рассмотрим краевую задачу (1)–(2) для уравнения колебаний струны, допускающую неклассический (неэйлеров) вариационный принцип с функционалом

$$D\,[u] = \int\limits_{\Omega} \left(u_x^2 + u_t^2 - 2u_x \cos x \sin t - 2u_t \sin x \cos t\right)\, dxdt + \int\limits_{\gamma_1} u^2 ds$$

С помощью нейронной сети, найдем приближенное решение u (x, t) этой краевой задачи, имеющей выход f (x, t; θ), где x, t – входные значения, а θ представляет собой вектор параметров (весов и смещений) нейронной сети.

$$u(x, t) \approx f(x, t; \theta).$$

Общая схема алгоритма

- 1. Выбрать начальные параметры нейронной сети θ_0 и скорость обучения α_0

- 2. Сгенерировать две случайные выборки:
  - n_0 точек из области Ω с распределением ν_0.
  - n_1 точек из границы $\Gamma_1 \subset \partial\Omega$ с распределением ν_1

- 3. Вычислить неэйлеров функционал $L_N(f)$ для сгенерированных случайных выборок, объединенных в мини-батч s_k:

$$\mathbf{s}_k = \left\{ \{(x_{i_0}, t_{i_0})\}_{i_0=1}^{n_0}, \{(x_{i_1}, t_{i_1})\}_{i_1=1}^{n_1} \right\},$$

$$L_0(f, \mathbf{s}_k; \boldsymbol{\theta}_k) = \frac{1}{n_0} \sum_{i_0=1}^{n_0} \left(f_x^2(x_{i_0}, t_{i_0}; \boldsymbol{\theta}_k) + f_t^2(x_{i_0}, t_{i_0}; \boldsymbol{\theta}_k) - \right.,$$

$$\left. -2 f_x(x_{i_0}, t_{i_0}; \boldsymbol{\theta}_k) \cos x_{i_0} \sin t_{i_0} - 2 f_t(x_{i_0}, t_{i_0}; \boldsymbol{\theta}_k) \sin x_{i_0} \cos t_{i_0} \right)^2,$$

$$L_1(f, \mathbf{s}_k; \boldsymbol{\theta}_k) = \frac{1}{n_1} \sum_{i_1=1}^{n_1} \left(f(x_{i_1}, t_{i_1}; \boldsymbol{\theta}_k) \right)^2,$$

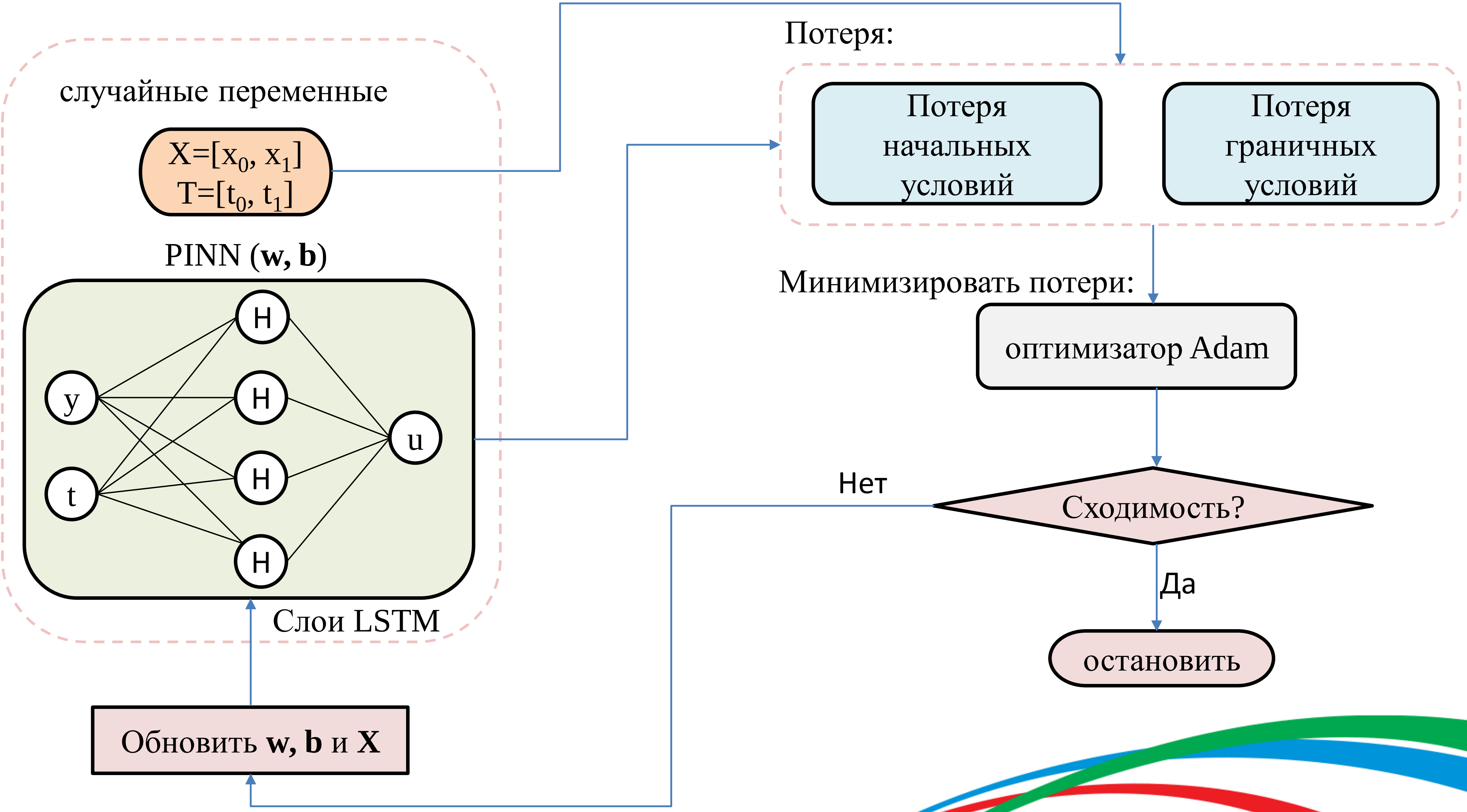
$$\mathcal{L}_N(f, \mathbf{s}_k; \boldsymbol{\theta}_k) = L_0(f, \mathbf{s}_k; \boldsymbol{\theta}_k) + L_1(f, \mathbf{s}_k; \boldsymbol{\theta}_k)$$

- 4. Выполнить шаги градиентного спуска с мини-батчем s_k и адаптивным алгоритмом (Adam):

$$\theta_{k+1} = \theta_k - \alpha_k \nabla_\theta L_N(f, s_k; \theta_k)$$

- 5. Повторять шаги 2-4 до тех пор, пока изменение параметров нейронной сети $\|\theta_{k+1} - \theta_k\|$ не станет достаточно малым.

Структура нейронной сети



фрагмент программы

model : Модель нейронной сети.
x0, t0: Выборочные точки в области Ω
x1, t1: Выборочные точки на границе Γ1

```
# Функция потерь
def loss(model, x0, t0, x1, t1):

  with tf.GradientTape() as tU0:
    U0 = model(x0, t0)

  U0x, U0t = tU0.gradient(U0, [x0, t0])

  L0 = tf.reduce_mean(tf.square(U0x) + tf.square(U0t) - \
                2.*U0x*tf.math.cos(x0)*tf.math.sin(t0) - \
                2.*U0t*tf.math.sin(x0)*tf.math.cos(t0))

  U1 = model(x1, t1)

  L1 = tf.reduce_mean(tf.square(U1))

  return L0 + L1, L0, L1
```

$$D[u] = \overbrace{\int_{\Omega} \left(u_x^2 + u_t^2 - 2u_x \cos x \sin t - 2u_t \sin x \cos t\right) dxdt}^{L0} + \underbrace{\int_{\gamma_1} u^2 ds}_{L1}$$

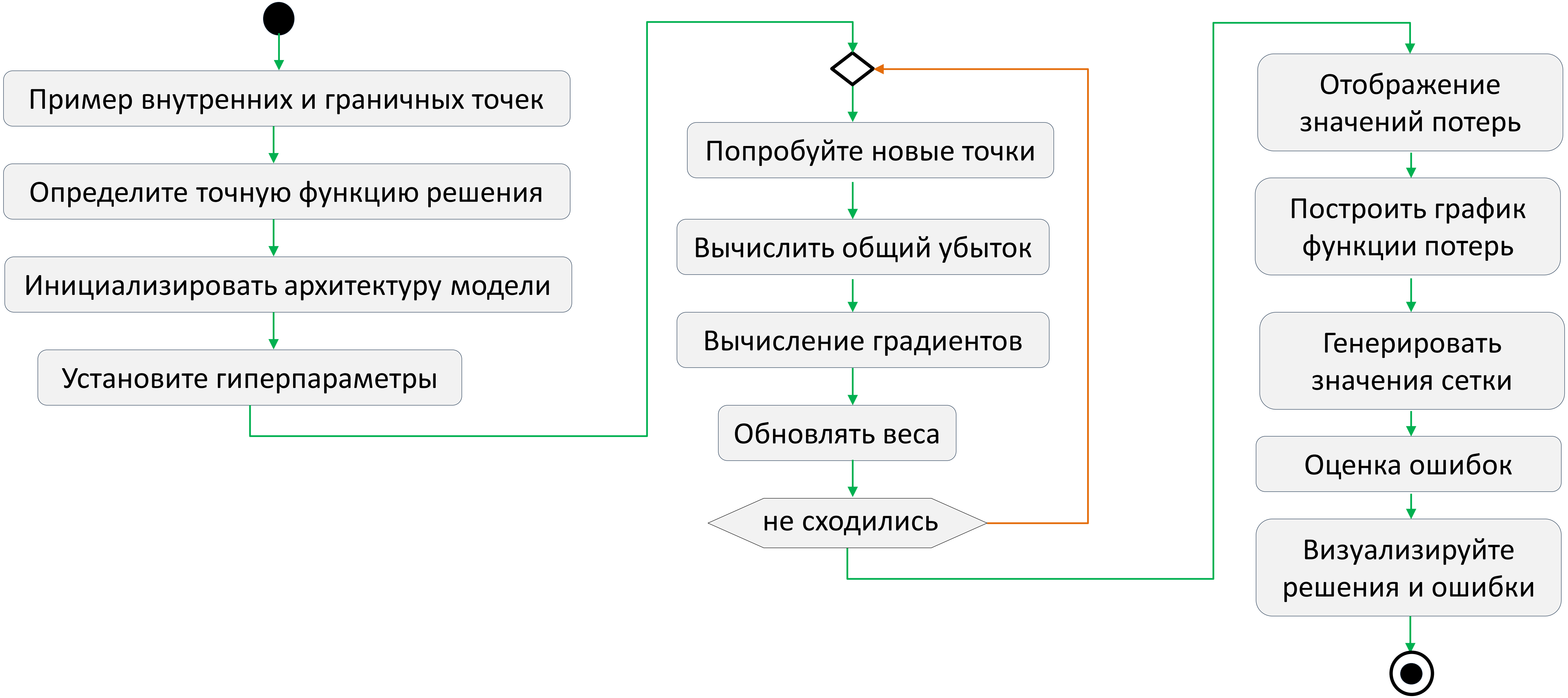
$\dfrac{\partial u_0}{\partial x_0}$ и $\dfrac{\partial u_0}{\partial t_0}$

функция вычисляет функцию потерь для обучения нейронной сети, объединяя два компонента:

- Потеря L_0, связанная с оператором дифференциального уравнения волны.
- Потеря L_1, связанная с начальными/граничными условиями.

Обе потери усредняются по выборочным точкам и комбинируются, чтобы сформировать общую функцию потерь, используемую для обучения модели нейронной сети.

Основные подходы к обучению



Визуализации Часть I

Рис. 1. На графике показана ошибка функции потерь при обучении LSTM на протяжении 100 этапов.

Визуализации Часть II

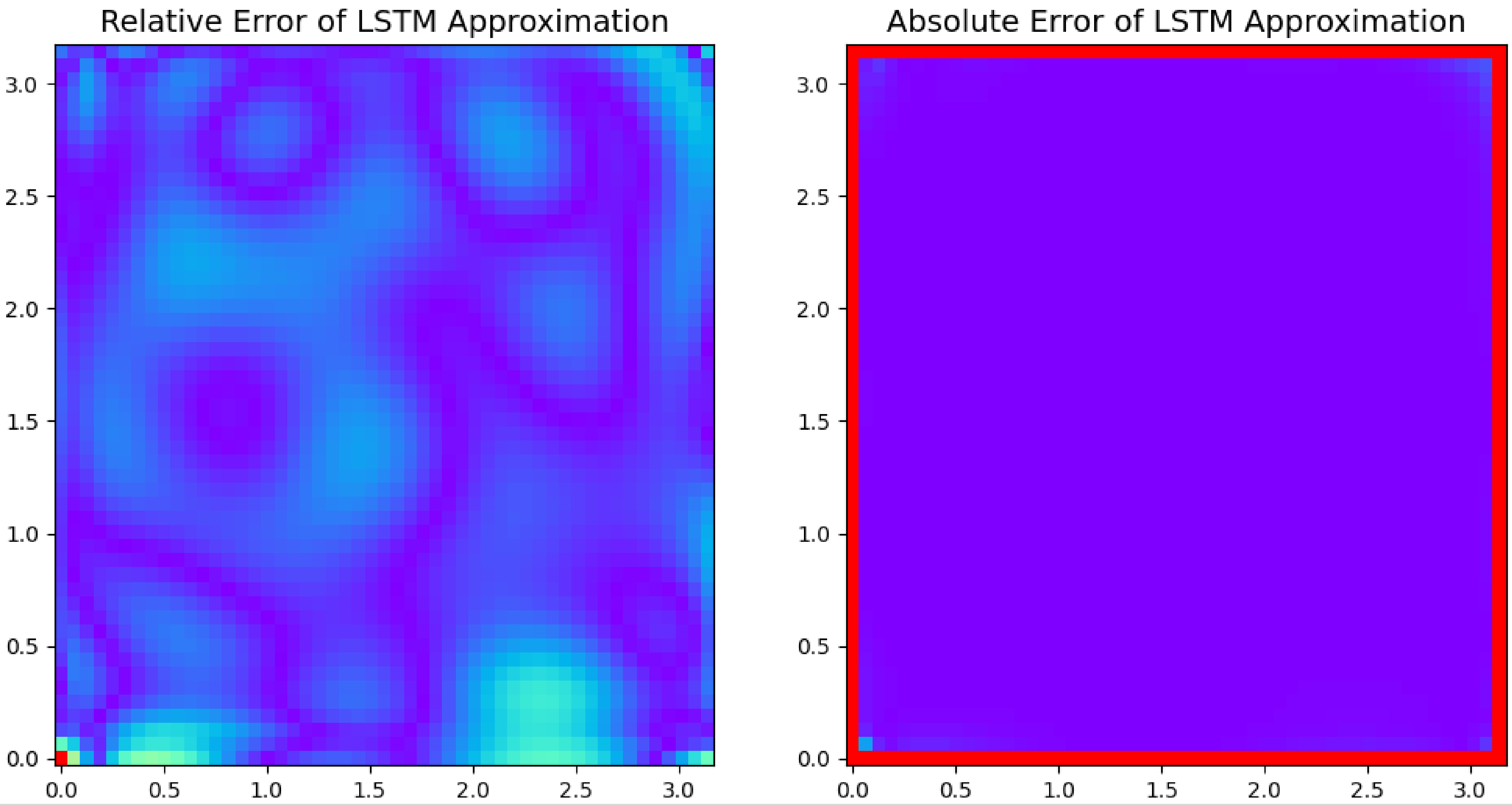


Рис. 2. График относительных ошибок показывает различия в точности модели во всей области, в то время как график абсолютных ошибок показывает, что ошибки, как правило, невелики.

Визуализации Часть III

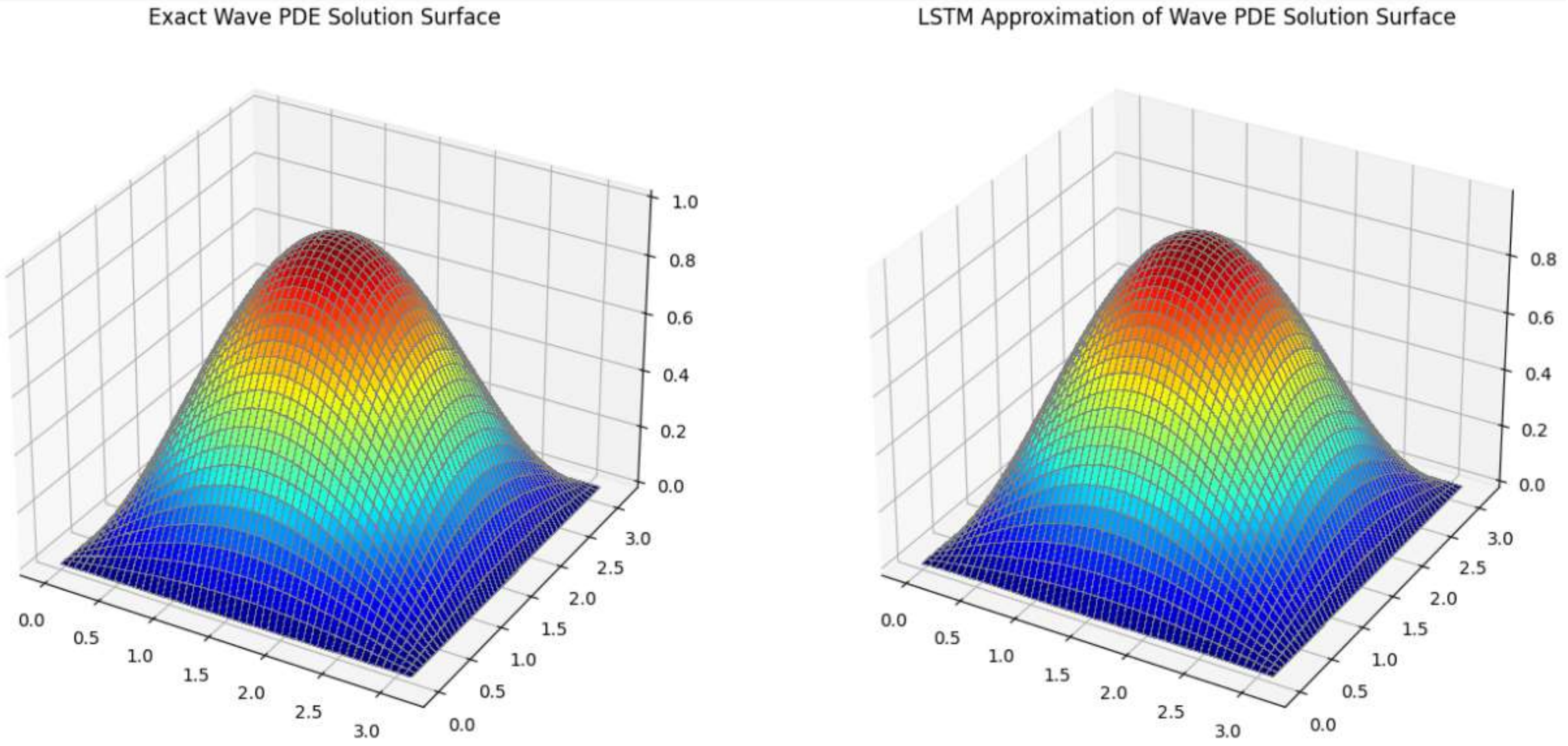


Рис. 3. эти два графика показывают, что предсказания модели LSTM очень похожи на точное решение, демонстрируя эффективность модели при изучении волнового (ДУЧП)

Заключение

В ходе выполнения выпускной работы:

- ✓ Проведен обзор дифференциальных уравнений, методов их решений и вариационных принципов.
- ✓ Изучено применение нейронных сетей для аппроксимации решений дифференциальных уравнений.
- ✓ Разработана неклассическая вариационная формулировка для одномерного волнового уравнения на основе метода В.М. Шалова.
- ✓ Преобразован вариационный функционал в форму, удобную для оценки методом Монте-Карло.
- ✓ Создана архитектура нейронной сети (PINN) для аппроксимации решения краевых задач.
- ✓ Реализованы и обучены нейронные сети для решения заданной краевой задачи.
- ✓ Проведены вычислительные эксперименты, которые подтвердили высокую точность модели.
- ✓ Произведена визуализация решений и их сравнение с точными решениями.

Перспективы дальнейших исследований включают совершенствование методов обучения нейронных сетей, расширение их применения на другие типы уравнений и задач, а также интеграцию дополнительных физических принципов в процесс обучения для повышения точности и эффективности вычислений.

БЛАГОДАРЮ ЗА ВНИМАНИЕ