

МИНОБРНАУКИ РОССИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМЕНИ Н.Г.ЧЕРНЫШЕВСКОГО»

Кафедра радиофизики и нелинейной динамики

**Применение методов машинного обучения
для классификации сложных сигналов**

АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ

студента 4 курса 4061 группы
направления 11.03.02 Инфокоммуникационные
технологии и системы связи
Института физики
Рябченко Андрея Дмитриевича

Научный руководитель

к.ф.-м.н., доцент

К.С. Сергеев

Зав. кафедрой радиофизики

и нелинейной динамики,

д.ф.-м.н., доцент

Г.И. Стрелкова

Саратов 2025 г.

ВВЕДЕНИЕ

Машинное обучение (machine learning, ML) — это класс методов искусственного интеллекта, характерной чертой которых является не прямое решение задачи, а обучение за счёт применения решений множества сходных задач. Суть машинного обучения заключается в создании алгоритмов и моделей, которые способны автоматически извлекать знания из данных и решать задачи или предсказывать результаты на их основе. Методы машинного обучения постоянно развиваются и находят все новые применения в различных отраслях.

Принято разделять методы машинного обучения на две группы: обучаемые с учителем и обучаемые без учителя. Приведём характерные примеры обоих классов.

1) Обучение с учителем:

Линейная регрессия. Моделирует линейную зависимость между входными и выходными данными.

Логистическая регрессия. Моделирует вероятность события на основе входных данных.

Дерево решений. Создает древовидную структуру для принятия решений на основе входных данных.

Метод опорных векторов. Находит гиперплоскость, которая разделяет данные на классы с максимальным запасом.

Нейронные сети. Иерархические структуры искусственных нейронов, которые обучаются выполнять сложные задачи.

2) Обучение без учителя:

Кластеризация. Группирует данные на основе сходств.

Авто-кодировщики. Нейронные сети, которые обучаются восстанавливать входные данные, улавливая их скрытые структуры.

Задачи решаемые с помощью методов машинного обучения также можно разделить на отдельные категории.

1) Задачи, включающие категориальный ответ, называются задачами классификации. Бинарной классификацией называют задачу разделения объектов на две категории или класса. Много-классовой классификацией называют задачу разделения объектов на более чем две категории или класса.

2) Задачи, в которых используется непрерывный отклик, называются задачами регрессии.

3) Задачи, в которых используется распределение объектов по категориям, когда неизвестно, сколько категорий получится в итоге, называются задачами кластеризации.

Целью данной выпускной квалификационной работы является построение методов классификации сложных сигналов с применением машинного обучения. Для достижения данной цели были поставлены следующие **задачи**:

1) Ознакомление с особенностями сложных сигналов:

а) Одиночных моделей радиофизических генераторов в хаотическом и квазипериодическом режиме

б) Цепочки связанных отображений с химерными структурами

2) Изучение различных методов машинного обучения

3) Применение машинного обучения для классификации хаотических и регулярных колебаний

4) Применение машинного обучения для обнаружения химерных состояний

5) Сопоставление различных методов машинного обучения и сравнение точности их работы

Выпускная квалификационная работа организована следующим образом. Первая часть посвящена исследованию классификации хаотических колебаний методами машинного обучения. Вторая часть посвящена исследованию

обнаружения химерных состояний методами машинного обучения. Для проведения всех исследований были написаны программы на языке программирования Python.

Основное содержание работы

1. Классификация хаотических и регулярных колебаний

Классические подходы обнаружения хаотической динамики требуют расчета множества характеристик и накладывают ограничение на минимальную длительность сигнала. С учётом этого, использование методов машинного обучения является перспективным подходом.

В данной работе исследовались следующие системы: генератор Анищенко-Астахова и генератор Рёсслера. Эти системы являются моделями демонстрирующими квазипериодические и хаотические колебания.

Генератор Анищенко-Астахова представляет собой трехмерную динамическую модель с тремя параметрами и описывается следующей системой дифференциальных уравнений:

$$\begin{aligned}\dot{x} &= mx + y - xz - dx^3 \\ \dot{y} &= -x \\ \dot{z} &= -gz + g\Phi(x)\end{aligned} \quad (1),$$

где первые два уравнения системы (1) представляют собой генератор Ван дер Поля, для которого $\dot{z}=0$ и $\Phi(x)=x^2$.

Генератор Рёсслера - трехмерная динамическая модель задаваемая следующей системой дифференциальных уравнений:

$$\begin{aligned}\dot{x} &= -y - z \\ \dot{y} &= x + ay \\ \dot{z} &= b + z(x - c)\end{aligned} \quad (2),$$

В данной работе предлагается метод обнаружения хаотического поведения системы на основе простых статистических характеристик с использованием методов машинного обучения. В качестве статистических характеристик в данном проекте использовались среднее значение, стандартное отклонение, асимметрия и эксцесс распределения.

В первую очередь исследовалась точность работы классификаторов на отдельных моделях (этап 1). Затем исследовалась точность работы классификаторов при обучении на одной модели а тестировании на другой (этап 2). В финале исследовалась точность работы классификаторов на объединённом наборе данных, состоящих из 2 моделей (этап 3).

	точность 0	F-мера 0	точность 1	F-мера 1
KNN (количество соседей равное 1)	0.928	0.896	0.892	0.904
SVM (показатель степени функции равный 10)	0.9	0.918	0.938	0.918
Нейронная сеть (количество эпох равное 6)	0.864	0.716	0.862	0.826

Таблица 1. Средняя точность работы классификаторов при их оптимальных параметрах, по результатам 2 и 3 этапа исследований. Основные параметры: «точность 0» и «F-мера 0» - точность определения и гармоническое среднее точности и полноты квазипериодики, «точность 1» и «F-мера 1» - точность определения и гармоническое среднее точности и полноты хаоса, «neighbors» - количество соседей, «degree» - показатель степени функции, «epochs» - количество эпох.

2. Классификация химерных состояний

В реальном мире все системы существуют взаимодействуя с другими, соответственно при исследовании реальных систем стоит говорить о наличии ансамблей связанных элементов. Поэтому одним из основных современных вопросов нелинейной динамики является взаимодействие элементов в ансамблях и установление сложных пространственно-временных структур. Новый расцвет нелинейной динамики пришел вместе с открытием таких пространственно-временных структур, как химерные состояния и уединенные состояния. В биологии химерами называются животные или растения, разные клетки которых содержат генетически разнородный материал в отличие от обычных организмов, у которых каждая клетка содержит один и тот же набор генов. Под термином «химерное состояние» («химера») в нелинейной динамике понимается такая пространственно-временная структура, когда в системе одновременно существует когерентный и некогерентный кластер. Уединенные состояния – это пространственно-временные паттерны, полученные в результате смещения нескольких узлов из основного синхронного кластера, которые обладают частотами, отличными от частот синхронного кластера. Эти состояния вызывают большой интерес для исследований. Такая форма динамики наблюдается в ансамблях нелокально взаимодействующих идентичных осцилляторов и заключается в появлении пространственно локализованных скоплений элементов с некогерентными (асинхронными) колебаниями, в то время как все остальные элементы ансамбля демонстрируют когерентную (синхронную) динамику. Химерные состояния были впервые описаны в системе связанных фазовых генераторов. С тех пор они также были обнаружены в широком спектре различных моделей, включая как отображения дискретного времени, так и непрерывные системы с регулярной и хаотической динамикой. Существование химерных состояний неоднократно подтверждалось не только численно, но и экспериментально. Однако, несмотря на большое количество работ, общие механизмы формирования химерных состояний в

различных системах остаются фактически неясными. Более того, до сих пор не существует классификации ансамблей взаимодействующих осцилляторов, позволяющей прогнозировать реализацию химерных состояний и их свойств в определенных классах систем. С учётом этого, использование методов машинного обучения является перспективным подходом.

В данной работе исследовалось кольцо нелокально связанных логистических отображений. Модель описывается следующим уравнением:

$$x_{n+1}^i = F(x_n^i) + \frac{\sigma}{2R} \sum_{j=i-r}^{i+r} [F(x_n^j) - F(x_n^i)] \quad (3),$$

где : x_{n+1}^i — парциальный элемент ансамбля, $i=1,2,\dots, N=1000$ — номер элемента, $F(x_n^i) = \lambda x_n(1-x_n)$, σ — сила связи, R — радиус связи.

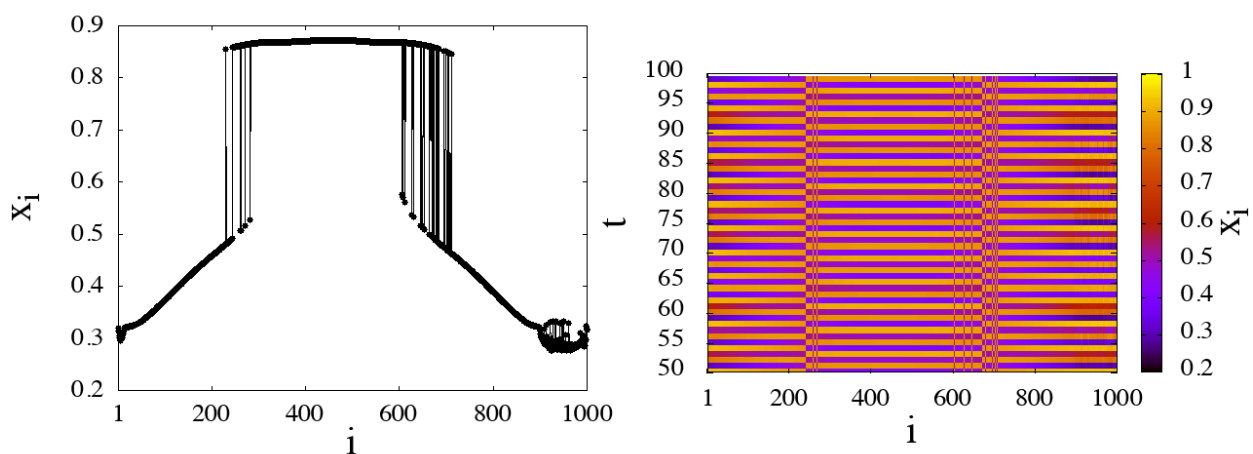


Рисунок 1. Мгновенный профиль (слева) и пространственно-временная диаграмма (справа) системы кольца нелокально связанных логистических отображений.

В данной работе предлагается метод обнаружения амплитудных химерных состояний на основе созданных временных реализаций, состоящих из значений координат парциальных элементов ансамбля в разные моменты времени, с использованием методов машинного обучения. По полученным результатам исследований можно заметить хорошую точность работы

классификаторов. В рамках исследования точности работы алгоритма искусственная нейронная сеть исследовались две структуры сети. Первая структура сети состояла из трёх слоёв: входного, скрытого и выходного. Исследовалась зависимость точности работы классификатора от количества нейронов в скрытом слое. По полученным результатам видно, что при достижении параметром n значения, равного шестнадцати, дальнейшее улучшение точности работы алгоритма не наблюдается. Данное значение является оптимальным для исследуемой структуры нейронной сети. Количество эпох равное двенадцати у алгоритма искусственная нейронная сеть в рамках исследований было признано оптимальным. Вторая структура сети состояла из двух слоёв: входного и выходного. Результаты также показали хорошую точность работы алгоритма. Количество эпох равное двенадцати у алгоритма искусственная нейронная сеть в рамках исследований было признано оптимальным.

	precision 0	f1-score 0	precision 1	f1-score 1
KNN (количество соседей равное 1)	1.00	1.00	1.00	1.00
SVM (показатель степени функции равный 10)	1.00	1.00	1.00	1.00
Искусственная нейронная сеть, с скрытым слоем (количество эпох равное 12, количество нейронов в скрытом слое равное 16)	1.00	0.99	0.97	0.99
Искусственная нейронная сеть, без скрытого слоя (количество эпох равное 12)	0.97	0.95	0.94	0.96

Таблица 2. Средняя точность работы классификаторов при их оптимальных параметрах. Основные параметры: «точность 1» и «F-мера 1» - точность определения и гармоническое среднее точности и полноты амплитудного химерного состояния, «точность 0» и «F-мера 0» - точность определения и гармоническое среднее точности и полноты всех остальных состояний, «neighbors» - количество соседей, «degree» - показатель степени функции, «epochs» - количество эпох, «n» - количество нейронов в скрытом слое.

Заключение

В рамках задачи применения машинного обучения для классификации хаотических и регулярных колебаний исследовалась точность работы классификаторов при различных условиях. В рамках 1 этапа исследовалась точность работы алгоритмов при обучении на отдельных системах. Были получены результаты, показывающие неплохую точность работы классификаторов (не ниже 75%). Лучше всего классификаторы определяют хаотическую динамику системы (точность не ниже 94%). Алгоритмы показывают высокую точность работы на генераторе Анищенко-Астахова. В рамках 2 этапа исследовалась точность работы классификаторов при обучении на одной модели, а тестировании на другой. Полученные результаты показали значительное снижение точности работы классификаторов. Это объясняется тем, что алгоритмы при обучении на одной из систем не разу не видели примеры из другой системы. При этом накопленного опыта при обучении на модели Анищенко-Астахова хватает с хорошей точностью классифицировать хаотическую динамику модели Рёсслера (не ниже 99%). А накопленного опыта при обучении на модели Рёсслера хватает с хорошей точностью классифицировать периодическую динамику модели Анищенко-Астахова. Из полученных результатов (таблица 1) можно сделать вывод, что лучшим классификатором в рамках задачи классификации хаотических и регулярных колебаний является SVM. Поэтому с его помощью можно обнаруживать хаотическое поведение сигналов с использованием методов машинного обучения на основе простых статистических характеристик. Таким образом, поставленная цель, а именно разработка алгоритма классификации хаотических и квазипериодических колебаний, была успешно достигнута.

В рамках задачи применения машинного обучения для обнаружения химерных состояний исследовалась точность работы классификаторов при различных условиях. Из полученных результатов (таблица 2) видно, что все использованные классификаторы показали хорошую точность работы (точность

не ниже 94%), но лучшими алгоритмами в рамках задачи поиска химерных состояний являются KNN и SVM. Таким образом, поставленная цель, а именно разработка алгоритма обнаружения амплитудных химерных состояний в системе кольца нелокальных связанных логистических отображений была успешно достигнута.