

Опыт построения и обучения нейро-нечеткой сети на основе байесовской логико-вероятностной модели

Г. А. Хамчичев¹, Г. И. Кожомбердиева²

Петербургский государственный университет путей сообщения Императора Александра I

¹gyqn@yandex.ru, ²kgi-liizht@yandex.ru

Аннотация. В докладе обсуждается опыт построения и обучения нейро-нечеткой сети на основе байесовской логико-вероятностной модели с использованием разработанного авторами инструментария. Рассматриваются примеры создания сетей, ориентированных на решение конкретных задач аппроксимации функций нескольких переменных. Представлены оптимистичные результаты сравнения с функциональностью инструмента ANFIS из пакета MATLAB, позволяющие говорить об эффективности предложенной модели нейро-нечеткой сети и возможностях ее применения в различных областях.

Ключевые слова: нейро-нечеткая сеть; байесовская логико-вероятностная модель нечеткого вывода; обучение нейро-нечеткой сети; аппроксимация функций; ANFIS

I. ВВЕДЕНИЕ

Настоящий доклад, который является продолжением доклада «О программной реализации нейро-нечеткой сети на основе байесовской логико-вероятностной модели» в данном сборнике, фокусируется на практическом применении разработанной программы. Задачей авторов доклада является демонстрация возможностей программы по созданию и обучению нейро-нечеткой сети (ННС) на основе байесовской логико-вероятностной модели нечеткого вывода (БЛВ-модели) на реальных и синтетических данных, что, с одной стороны, позволяет оценить аппроксимационные способности такой сети и эффективность ее обучения, а с другой – сравнить разработанную программу с инструментом ANFIS из MATLAB. В докладе рассматриваются примеры создания сетей, ориентированных на решение конкретных задач аппроксимации функций нескольких переменных.

II. ПОРЯДОК РАБОТЫ С ПРОГРАММОЙ, РЕАЛИЗУЮЩЕЙ ННС НА ОСНОВЕ БЛВ-МОДЕЛИ, ПРИ ПОСТРОЕНИИ И ОБУЧЕНИИ СЕТИ

Для загрузки наборов данных используются элементы управления, расположенные на панели «Загрузка данных» окна программы. Программа поддерживает загрузку наборов трех типов: обучающий, валидационный и тестовый. Для создания сети пользователь указывает обучающий набор данных, количество термов входных ЛП и начальное значение весов w . На этой основе программа автоматически создает нечеткие правила, необходимые для

формирования структуры ННС, и определяет начальные значения параметров сети c , σ , \tilde{y} .

Для начала обучения нужно выбрать алгоритм обучения, ввести целевое значение ошибки и количество эпох обучения сети. Обучение созданной ННС начинается по нажатию кнопки «Старт». Пользователь может в любой момент приостановить обучение сети, продолжить этот процесс, и завершить его.

В программе реализована возможность настройки параметров обучения, позволяющая отключать и включать обновление параметров для каждого параметрического слоя ННС, менять шаги их изменения, выбирать направление градиента и стратегию изменения значений параметров обучения (шагов изменения параметров сети) во время обучения, а также вручную настраивать параметры стратегий обучения.

На данный момент в программе поддерживаются две стратегии изменения значений параметров обучения. Одна из них – классическая стратегия, используемая в MATLAB и описанная в [1]. Суть ее состоит в том, что при уменьшении ошибки на протяжении четырех эпох значение шага увеличивается, а если ошибка претерпевает две последовательные комбинации из одного увеличения и одного уменьшения – шаг уменьшается. Второй стратегией является вариант с возвратом значений параметров на определенное задаваемое число эпох назад при увеличении ошибки.

После завершения обучения сеть подлежит проверке с использованием наборов данных для валидации и тестирования, что позволяет оценить ее способность к обобщению и точность предсказаний на новых данных. Валидационный набор данных, который не участвует в обучении сети, позволяет настраивать параметры обучения и предотвращает ее переобучение. Тестовый набор, аналогично валидационному, состоит из данных, не участвующих в обучении, и предназначен для окончательной проверки обобщающей способности обученной сети после тонкой настройки и валидации, имитируя ее реальное применение [2]. Для проверки сети на панели «Проверка сети» нужно выбрать предварительно загруженный набор данных и нажать кнопку «Начать проверку». Если набор данных корректен, то результаты, включая значения ключевых метрик, отобразятся в специализированном текстовом поле графического интерфейса программы.

Пользователь может проверять работу сети на собственных данных, вводя значения входных переменных для получения дефазифицированного результата. Это позволяет проводить непосредственную проверку эффективности обученной ННС на специфичных для конкретного пользователя данных, упрощая процесс ее тестирования и адаптации под индивидуальные нужды. Результат расчета на основе текущих значений параметров сети отображается в специальном текстовом поле, включающим также и оценку ошибки вычисления в виде границ доверительного интервала \pm RMSE (Root Mean Squared Error, корень из среднеквадратической ошибки).

Пользователь имеет возможность сохранить текущее состояние сети со всеми значениями параметров, чтобы в дальнейшем можно было к нему вернуться. При экспорте и выборе опции «Совместимость» сеть будет экспортирована в формат, совместимый с авторской программой для решения задач нечеткого вывода на основе байесовской логико-вероятностной модели, зарегистрированной в Роспатенте [3].

III. ПРИМЕРЫ СОЗДАНИЯ СЕТЕЙ, ОРИЕНТИРОВАННЫХ НА РЕШЕНИЕ ПРАКТИЧЕСКИХ ЗАДАЧ

Для исследования возможностей предложенной модели ННС выбрано несколько наборов данных: два достаточно известных [4, 5], доступных на различных ресурсах, и два синтетических. Наборы [4, 5] являются хорошо зарекомендовавшими себя открытыми наборами. Они доступны на ресурсах Delve Datasets и UCI Machine Learning Repository, представляющих собой обширные коллекции репозитория для исследователей в области машинного обучения и анализа данных. Эти наборы доступны также на платформе Kaggle [6], предоставляющей возможности для участия в соревнованиях и доступ к обширной базе данных для исследований.

Наборы данных [4, 5] имеют большую размерность, из-за чего их обработка имеет некоторые ограничения даже в профессиональном инструменте MATLAB, поскольку при попытке создать сеть на основе целого набора данных возникает нехватка памяти. Связано это с тем, что сложность сети экспоненциально зависит от количества термов лингвистических переменных (ЛП), построенных на основе входных переменных. Так как в большинстве случаев входные переменные в разной мере влияют на результат, было принято решение о сокращении размерности этих наборов данных до трех входных переменных и одной выходной. Для этого была использована техника сокращения размерности пространства признаков на основе корреляционного анализа взаимосвязи между каждым из признаков и целевой переменной. Кроме того, была произведена очистка входных данных путем удаления строк с отсутствующими значениями.

В докладе сравниваются ННС на основе БЛВ-модели и ННС типа ANFIS с одинаковыми значениями количества термов входных ЛП. В качестве функций принадлежности (ФП) используются гауссовы ФП. Используемым алгоритмом обучения является гибридный алгоритм.

Для сравнения с ANFIS используется метрика качества RMSE, поскольку, с одной стороны, она адекватно позволяет оценить шум в данных, а с другой это единственная метрика, доступная как в MATLAB, так и в разработанной авторами программе.

Для демонстрации процесса обучения сетей и его сходимости, представлены графики зависимости RMSE от числа использованных эпох обучения. Поскольку сети демонстрируют разные темпы сходимости, графики ограничены начальным периодом обучения, когда наблюдается активное уменьшение ошибки. Полная сходимость каждой модели, достигнутая в последующих эпохах, характеризуется минимальными изменениями RMSE и представлена в виде итоговых значений в соответствующих таблицах.

A. Пример Boston Housing

Набор данных о ценах на жилье в Бостоне ВН [4] содержит информацию о различных характеристиках жилых домов в окрестностях Бостона, включая среднее количество комнат в доме, уровень преступности, налоги на недвижимость и многое другое (всего в наборе используется 14 характеристик). На основе корреляционного анализа выделено три наиболее важных фактора, влияющих на среднюю стоимость домов *medv*, а именно *lstat* (-0,74), *rm* (0,70) и *ptratio* (-0,51).

После удаления строк с пустыми значениями набор содержит 452 строк. Из них 271 строка выделена для формирования обучающего набора данных, 90 – для валидационного набора и 91 – для тестового набора.

График на рис. 1 иллюстрирует зависимость RMSE от числа эпох обучения для первых 50 эпох.

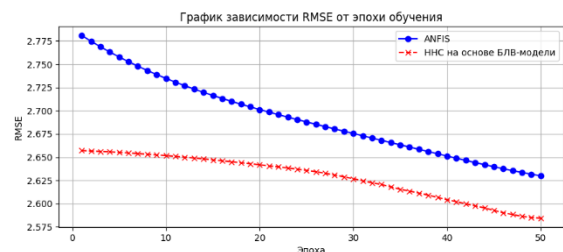


Рис. 1. График зависимости RMSE от эпохи обучения для набора ВН на первых 50 эпохах

ТАБЛИЦА I. СРАВНЕНИЕ ННС ПО МЕТРИКЕ RMSE ДЛЯ НАБОРА ВН

Сеть	ННС на основе БЛВ-модели			ANFIS		
	обуч.	вал.	тест.	обуч.	вал.	тест.
Набор данных						
До обучения	7,41	6,15	13,86	26,53	25,77	20,81
Лучшая эпоха	2,58	2,86	10,20	2,61	2,67	7,47

Анализируя данные рис. 1 и табл. I, можно отметить значительное улучшение показателя RMSE после обучения как для ННС на основе БЛВ-модели, так и для ANFIS. График подтверждает их близкую эффективность и сходный темп обучения. Однако, несмотря на сходные результаты на этапе обучения и валидации, на тестовом наборе данных ННС на основе БЛВ-модели демонстрирует RMSE 10,2 против 7,47 у

ANFIS, что указывает на более высокую способность ANFIS к обобщению на новых данных. Тем не менее, близкие значения показателей на обучающем и валидационном наборах подчеркивают большой потенциал ННС на основе БЛВ-модели в качестве инструмента аппроксимации сложных функциональных зависимостей, особенно учитывая, что ANFIS требует доступа к платному программному обеспечению MATLAB.

В. Пример Auto MPG

Набор данных Auto MPG [5] описывает различные характеристики автомобилей и содержит 9 переменных. С помощью корреляционного анализа их количество было сокращено до 3: *cylinders* (-0,775), *displacement* (-0,804) и *mpg*. В наборе 398 строк: 238 из них были отведены на обучение, 79 – на валидацию и 81 – на тест. Результаты обучения приведены на рис. 2 и в табл. II.

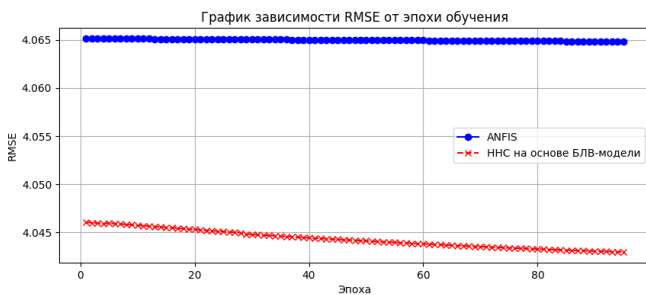


Рис. 2. График зависимости RMSE от эпохи обучения для набора Auto MPG

ТАБЛИЦА II. СРАВНЕНИЕ ННС ПО МЕТРИКЕ RMSE ДЛЯ НАБОРА AUTO MPG

Сеть	ННС на основе БЛВ-модели			ANFIS		
	обуч.	вал.	тест.	обуч.	вал.	тест.
Набор данных						
До обучения	13,915	14,234	13,428	24,95	24,426	24,609
Лучшая эпоха	4,04	4,2844	4,1762	4,0648	5,0002	4,2639

Результаты показывают, что разница значений RMSE у обеих моделей ННС невелика как на начальном этапе, так и после обучения, что свидетельствует о сравнимом качестве аппроксимации данных обеими моделями. При этом, в отличие от ANFIS, ННС на основе БЛВ-модели демонстрирует тенденцию уменьшения значения этого показателя в ходе обучения.

С. Пример набора данных на основе нелинейной функции одной переменной с добавлением шума.

Синтетический набор данных СНД1 сгенерирован на основе функции $f(x) = -4x^3 + 8x^2 + 5x - 1$. Диапазон значений входной переменной: $[-1,5; 3]$. К результатам вычисления $f(x)$ добавлен гауссов шум с распределением $N(0; 1)$ для имитации реальных условий измерений.

Изначально было сгенерировано 250 строк, после чего были удалены строки с неуникальными значениями аргумента. Итоговое количество строк составило 200. Из них 120 строк выделены для формирования обучающего набора данных и по 40 – для валидационного

и тестового наборов. Результаты обучения сетей приведены на рис. 3 и в табл. III.

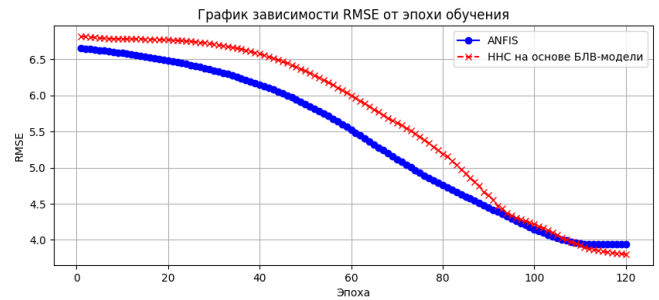


Рис. 3. График зависимости RMSE от эпохи обучения для СНД1 на первых 120 эпохах

ТАБЛИЦА III. СРАВНЕНИЕ ННС ПО МЕТРИКЕ RMSE ДЛЯ СНД1

Сеть	ННС на основе БЛВ-модели			ANFIS		
	обуч.	вал.	тест.	обуч.	вал.	тест.
Набор данных						
До обучения	9,914	12,569	9,367	8,238	10,238	7,599
Лучшая эпоха	3,726	4,986	4	3,941	4,313	3,758

Анализ результатов показывает, что обе модели ННС достигли значительного уменьшения RMSE по итогам обучения. График зависимости RMSE от числа эпох обучения демонстрирует стабильное улучшение качества обеих моделей. По завершении этапа обучения ННС на основе БЛВ-модели достигает более высокого качества, чем ANFIS. Согласно данным табл. III, ННС на основе БЛВ-модели показывает лучшие результаты на обучающем наборе данных, что подтверждает ее высокую адаптивность и способность к аппроксимации нелинейных зависимостей в условиях шума.

Д. Пример набора данных на основе нелинейной функции трех переменных с добавлением шума.

Синтетический набор данных СНД3 сгенерирован на основе функции трех переменных $f(x, y, z) = x^2 - 4y + 5z + 2$. К значениям функции $f(x, y, z)$ добавлен гауссов шум с распределением $N(0; 0,5)$ для имитации реальных условий измерений. Диапазон значений входных переменных $[-4; 4]$.

Набор данных состоит из 1000 строк, 600 строк из которых выделены для формирования обучающего набора данных, 200 – для валидационного набора и оставшиеся 200 – для тестового набора. Результаты обучения сетей приведены на рис. 4 и в табл. IV.

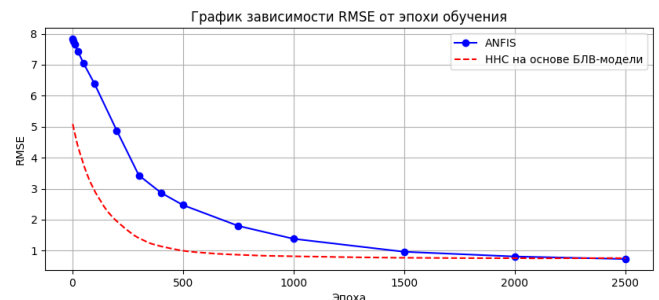


Рис. 4. График зависимости RMSE от эпохи обучения для СНД3 на первых 2500 эпохах

ТАБЛИЦА IV. СРАВНЕНИЕ ННС ПО МЕТРИКЕ RMSE ДЛЯ СНДЗ

Сеть	ННС на основе БЛВ-модели			ANFIS		
	обуч.	вал.	тест.	обуч.	вал.	тест.
Набор данных						
До обучения	70,33	68,958	70,663	61,783	62,496	59,67
Лучшая эпоха	0,7153	0,7908	0,7220	0,5237	0,5203	0,5426

Было проведено множество экспериментов с обучением ННС на основе БЛВ-модели и подбора наилучших значений параметров обучения. Результат с наименьшей полученной ошибкой представлен в таблице IV. Дополнительная информация к этому результату: лимит эпох – 2500, начальные шаги для весов правил и параметров ФП – 0,00005 и 0,0005 соответственно, количество термов входных ЛП – 3, начальные веса правил равны 0,5. Параметры, установленные для обучения сети ANFIS (количество термов входных ЛП – 3, количество эпох – 10000) при этом не менялись.

График на рис. 4 показывает, что ННС на основе БЛВ-модели достигает стабильно низких значений ошибки быстрее, чем ANFIS, что свидетельствует в пользу предложенного подхода.

Из табл. IV видно, что, несмотря на снижение ошибки RMSE по сравнению с начальными значениями для обеих моделей, сеть ANFIS показывает лучшие результаты на тестовом наборе данных по сравнению с ННС на основе БЛВ-модели. Однако достигнутые результаты все же говорят и о высоком потенциале ННС на основе БЛВ-модели как универсального аппроксиматора сложных функциональных зависимостей.

В настоящее время для ННС на основе БЛВ-модели при работе с СНДЗ не удалось получить значение ошибки RMSE менее 0,7, однако авторы считают, что это возможно при подборе корректной стратегии обучения, поскольку сеть имеет больше параметризованных слоев и поэтому может быть более гибко настроена. Однако это связано с возрастающей сложностью обучения и необходимостью более тщательного подбора значений таких параметров обучения, как шаги изменения параметров сети. Корректная настройка этих параметров критична для обеспечения сходимости сети, что требует дополнительных исследований и экспериментов для достижения наилучших результатов.

IV. ЗАКЛЮЧЕНИЕ

Результаты, полученные в процессе работы с программой для построения и обучения ННС на основе БЛВ-модели нечеткого вывода, являются основанием для оптимистичных выводов о способности ННС этого типа к обучению и решению практических задач в областях, в которых обычно применяются ННС, в частности ANFIS, а также о сопоставимой эффективности этих сетей.

Эксперименты подтвердили эффективность применения гибридного алгоритма обучения. Сети, обученные с помощью этого алгоритма (примеры обучения приведены в докладе), показывают лучшие результаты аппроксимации, чем сети, обученные с

помощью алгоритма обратного распространения ошибки.

В ходе экспериментов было выявлено, что шаги изменения нелинейных параметров сети (весов нечетких правил и ФП входных ЛП) должны быть сравнимы друг с другом и не превышать 0,01, а шаг для линейных параметров должен изменяться в пределах 0,1 – 1, чтобы обновление параметров сети выполнялось согласованно, и ошибка постоянно уменьшалась в процессе обучения. Следует продолжить исследование способов согласования параметров обучения, в том числе использование стратегий с асинхронным изменением шагов обучения.

В дополнение к сказанному, авторы считают необходимым особо выделить следующее.

В докладах на конференции SCM'2023 [7, 8] отмечалось, что объединение в гибридных системах нейросетевого и нечеткого подходов позволяет, с одной стороны, привнести способность к обучению и параллелизм вычислений, присущие нейронным сетям, в системы нечеткого вывода, а с другой стороны – усилить интеллектуальные возможности нейронных сетей лингвистически интерпретируемыми, нечеткими правилами выработки решений.

По мнению авторов, в программе для построения и обучения ННС на основе БЛВ-модели нечеткого вывода, создание кода которой на языке Java обеспечил Г.А. Хамчиев, аппарат нечеткого вывода теряет свою семантическую природу. “Нечеткая” терминология лишается смысловой значимости, нечеткие правила упрощаются и приобретают скорее формальный характер, ЛП, термы ЛП, ФП используются просто как элементы модели нейронной сети.

Нейросетевая структура, поддерживаемая адаптированными версиями алгоритмов обучения, в разработанной программе для построения и обучения ННС становится самодовлеющей.

СПИСОК ЛИТЕРАТУРЫ

- [1] Neuro-Adaptive Learning and ANFIS [Электронный ресурс]. URL: <https://www.mathworks.com/help/fuzzy/neuro-adaptive-learning-and-anfis.html> (дата обращения 01.03.2024).
- [2] Бишоп К.М. Распознавание образов и машинное обучение / Пер. с англ. СПб.: ООО Диалектика. 2020. 960 с.
- [3] Кожомбердиева Г.И., Хамчиев Г.А., Бураков Д.П. Программа для решения задач нечеткого вывода на основе байесовской логико-вероятностной модели. Свидетельство о государственной регистрации программы для ЭВМ RU 2021662943 от 09.08.2021.
- [4] The Boston Housing Dataset [Электронный ресурс]. URL: <https://www.cs.toronto.edu/~delve/data/boston/bostonDetail.html> (дата обращения 27.03.2024).
- [5] Auto MPG [Электронный ресурс]. URL: <https://archive.ics.uci.edu/dataset/9/auto+mpg> (дата обращения 27.03.2024).
- [6] Kaggle [Электронный ресурс]. URL: <https://www.kaggle.com/> (дата обращения 09.04.2024).
- [7] Кожомбердиева Г.И., Бураков Д.П., Хамчиев Г.А. О нейросетевой интерпретации байесовской логико-вероятностной модели нечеткого вывода // Международная конференция по мягким вычислениям и измерениям. 2023. Т. 1. С. 25–28.
- [8] Хамчиев Г.А., Кожомбердиева Г.И. О настройке и возможностях обучения нейро-нечеткой сети на основе байесовской логико-вероятностной модели // Международная конференция по мягким вычислениям и измерениям. 2023. Т. 1. С. 29–33.