

Федеральное государственное бюджетное образовательное учреждение
высшего профессионального образования
«Тольяттинский государственный университет»

Институт энергетики и электротехники

Кафедра «Электроснабжение и электротехника»

Направление подготовки магистра
13.04.02 «Электроэнергетика и электротехника»

Магистерская программа «Режимы работы электрических источников
питания, подстанций, сетей и систем»

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ

на тему «Повышение эффективности работы энергосбытовой компании на
оптовом рынке электроэнергии и мощности»

Студент(ка)	<u>Р.В. Рузанов</u> (И.О. Фамилия)	_____	(личная подпись)
Руководитель	<u>Д.А. Кретов</u> (И.О. Фамилия)	_____	(личная подпись)
Консультанты	_____	_____	(личная подпись)
	_____	_____	(личная подпись)

Руководитель магистерской
программы д.т.н., профессор В.В. Вахнина
(ученая степень, звание, И.О. Фамилия) _____
(личная подпись)
« _____ » _____ 2016 г.

Допустить к защите

Заведующий кафедрой д.т.н., профессор В.В. Вахнина
(ученая степень, звание, И.О. Фамилия) _____
(личная подпись)
« _____ » _____ 2016 г.

Тольятти 2016

Содержание

Введение	4
1 Анализ и сравнение существующих моделей прогнозирования	8
1.1 Метод экспертной оценки	8
1.2 Статистические методы прогнозирования	13
1.3 Методы прогнозирования, основанные на использовании искусственных нейронных сетей	20
1.4 Разомкнутые искусственные нейронные сети	29
1.5 Многослойные нейронные сети с перекрестными связями	32
1.6 Преимущества нейронных сетей	33
1.7 Сравнение принципов работы человеческого мозга и искусственных нейронных сетей	40
1.8 Модели нейронов	45
1.9 Принципы обучения нейронных сетей	48
1.10 Обучение, основанное на коррекции ошибок	50
1.11 Обучение на основе памяти	54
1.12 Парадигмы обучения нейронных сетей	56
1.13 Задачи обучения	61
1.14 Обучающие данные для нейронной сети	64
Вывод по главе 1	69
2 Разработка модели прогнозирования на базе искусственной нейронной сети	70
2.1 Выбор программного продукта для реализации модели прогнозирования на базе искусственной нейронной сети	70
2.2 Составление базы данных почасового электропотребления	72
2.3 Адаптация входных данных для обучения искусственной нейронной сети	73
2.4 Выбор алгоритма обучения нейронной сети	74
2.5 Выбор архитектуры нейронной сети	74

2.6	Сравнение результатов прогнозирования при различных моделях построения ИНС	78
	Выводы по главе 2	84
3	Введение в модель прогнозирования сведений о метеофакторах, влияющих на формирование графика электропотребления	85
3.1	Оценка эффективности прогнозирования выбранной модели нейронной сети на различных типах данных	85
3.2	Оценка эффективности введения в прогнозную модель возможности учета погодных данных	87
3.3	Оценка эффективности прогнозирования выбранной модели нейронной сети с возможностью учета метеофакторов	89
	Выводы по главе 3	92
	Заключение	93
	Список использованных источников	95

ВВЕДЕНИЕ

Электроэнергетика – это системообразующая отрасль промышленности Российской Федерации от успешного развития которой зависит в первую очередь энергобезопасность страны, а также интенсивность роста всех отраслей экономики государства. Электроэнергетику можно разделить на 4 основные сегмента: генерация, передача, распределение и сбыт. Одной из особенностей электроэнергии, как товара является невозможность ее хранения и накопления, поэтому ее потребления происходит одновременно с выработкой, что в свою очередь обуславливает особенности торговли таким видом товара.

Для производства электроэнергии генерирующим компаниям необходимо поддерживать в постоянной готовности генерирующие мощности, а также иметь запасы топлива в количестве, необходимом для удовлетворения спроса на электроэнергию в энергосистеме страны в данный конкретный момент времени. По такому же принципу работают и сетевые организации, функцией которых является доставка электроэнергии до конечного потребителя. Стоимость их услуг также входит в конечную стоимость каждого киловатта потребленной электроэнергии.

В связи с этим, можно сделать вывод о том, что для обеспечения бесперебойной работы всей цепочки предприятий, начиная с генерации и заканчивая конечным потребителем, необходима предварительная оплата всех затрат, связанных с работой этих предприятий для гарантированной поставки электроэнергии в необходимом объеме и качество которой соответствовало бы ГОСТ 32144-2013.

В свою очередь, потребителю электроэнергии для планирования своих затрат по этой статье расходов нужно спрогнозировать объемы электроэнергии, необходимой для нормального функционирования своего производства. В такой ситуации перед потребителем встает вопрос: каким образом это сделать и кто сможет это сделать? На данный момент

существует два варианта решения этой проблемы: самостоятельный выход предприятия на оптовый рынок электроэнергии (мощности) (ОРЭМ) и заключение договора энергоснабжения с энергосбытовой компанией или гарантирующим поставщиком.

В случае самостоятельного выхода на ОРЭМ, у предприятия появляется возможность покупать электроэнергию по цене ниже розничной, но в то же время возникают обязательства по ежедневному почасовому планированию электропотребления, созданию собственной автоматизированной информационно-измерительной системы коммерческого учета электроэнергии (АИИС КУЭ), которая отвечала бы техническим требованиям ОРЭМ. Также существует ограничение по суммарной присоединенной мощности энергопринимающего оборудования, которая составляет не менее 20 МВА суммарно по предприятию и в каждой группе точек поставки в отдельности составляет не менее 750 кВА. Поэтому для многих предприятий наиболее простым и наименее затратным является вариант заключения договора энергоснабжения с энергосбытовой компанией или гарантирующим поставщиком.

В таком случае, все обязанности по планированию электропотребления и покупке электроэнергии на ОРЭМ ложатся на энергосбытовую компанию, а предприятие в свою очередь покупает электроэнергию по розничной цене. Стоит также учесть, что клиентами энергосбытовых компаний, как правило являются большое множество таких предприятий, графики электропотребления которых абсолютно различны, а в случае с гарантирующими поставщиками в обязательном порядке особой категорией потребителей таких компаний является население, которому согласно Постановлению Правительства РФ №442 от 04.05.12г. «О функционировании розничных рынков электрической энергии, полном и (или) частичном ограничении режима потребления электрической энергии» гарантирующий поставщик обязан оказывать все необходимые услуги по электроснабжению. Стоит также отметить, что согласно тому же нормативному документу,

гарантирующий поставщик обязан быть участником торгов на ОРЭМ. Сложная структура электропотребления, включающая в себя предприятия с различными присоединенной мощностью и графиками работы, а также население и приравненные к нему категории потребителей, значительно осложняют прогнозирование для энергосбытовых компаний и гарантирующих поставщиков.

Стоит также отдельно сказать об особенностях потребления электроэнергии населением. Специфика его такова, что оно в значительной степени зависит от погодных условий, естественного освещения и статуса дня (рабочий или выходной день). Поэтому графики электропотребления имеют значительные отличия в различные дни недели, времена года, а также в значительной степени зависят от включения централизованного отопления в жилых домах.

Также качественное планирование электропотребления приводит к снижению затрат энергосбытовых компаний на ОРЭ(М), что в свою очередь позволяет ей экономить средства и перенаправлять их на развитие компании.

В целом такое устройство рынка электроэнергии имеет своей конечной целью рациональное использование энергетических и экономических ресурсов. Эти принципы описаны в Федеральном Законе РФ № 261-ФЗ от 23 ноября 2009 г. «Об энергосбережении и о повышении энергетической эффективности». Вследствие повышения цен на энергоресурсы в данный момент энергосбережение становится приоритетным направлением при снижении затрат на создание конечной продукции.

Поэтому задача создания моделей прогнозирования электропотребления энергосбытовых компаний является актуальной для разработки в научном сообществе.

На данный момент существует несколько способов составления прогнозных графиков электропотребления. Одним из наиболее популярных является метод экспертной оценки. Он заключается в том, что составление графиком электропотребления зависит от опыта и интуиции эксперта. В

качестве эксперта в таких случаях выступает инженер-энергетик или ведущий инженер предприятия. Популярность данного метода обусловлена простотой его реализации и отсутствием необходимости разрабатывать или же приобретать дорогостоящее программное обеспечение для прогнозирования.

Однако, в виду существенных недостатков такого метода, таких как высокая зависимость от субъективного мнения эксперта-аналитика и невозможность качественного составления прогноза при отсутствии такого сотрудника на рабочем месте, многие предприятия начинают разработку собственных моделей прогнозирования.

Это предполагает большой объем работы по сбору статистических данных, а также выбору моделей прогнозирования.

Поэтому целью магистерской диссертации является повышение эффективности работы энергосбытовой компании на оптовом рынке электроэнергии (мощности) в результате разработки модели прогнозирования электропотребления на основе искусственных нейронных сетей.

Для достижения поставленной цели необходимо решить следующие основные задачи:

- анализ и сравнение существующих методов прогнозирования;
- выбор программного продукта, на основе которого будет происходить построение модели прогнозирования;
- анализ и сравнение существующих моделей искусственных нейронных сетей, их архитектуры и способов обучения;
- испытание выбранной модели на тестовой выборке данных;
- расчет экономического эффекта от использования полученной модели прогнозирования в работе энергосбытовой компании.

1 Анализ и сравнение существующих моделей прогнозирования

1.1 Метод экспертной оценки

Прогнозирование объемов почасового потребления электроэнергии – серьезная проблема, с которой сталкиваются предприятия, работающие на оптовом рынке электроэнергии и мощности. В условиях функционирования ОРЭМ для энергосбытовых компаний важно с максимальной точностью прогнозировать объемы потребления электроэнергии. В противном случае, компания несет дополнительные финансовые затраты, перекладываемые на конечного потребителя, теряя за счет этого свою привлекательность на рынке. В данной ситуации необходим прогноз "на сутки вперед" в разрезе часовых (получасовых) интервалов. Точность напрямую зависит от методик расчета, позволяющих минимизировать покупку/продажу энергии в рамках балансирующего рынка.

Необходимость точного прогнозирования значений электропотребления в последние годы стала особенно острой. Главным образом это вызвано реформированием электроэнергетической отрасли и созданием конкурентного рынка электроэнергии и мощности, в результате чего от точности прогнозирования зависит финансовое благополучие предприятий. В России процессы либерализации электроэнергетики несколько отстают от западных стран. Поэтому у нас существует дефицит современных методик прогнозирования, специализированных программных продуктов для работы на ОРЭМ, а также специалистов в данной области. Проблема осложняется еще и тем, что предприятия не готовы делиться собственными методиками прогнозирования, которые приносят им прибыль в условиях рыночной конкуренции.

В настоящее время в России на промышленных предприятиях все еще распространен "ручной прогноз", основанный на личном опыте и интуиции эксперта: специалиста отдела, главного энергетика. В общем же случае,

экспертом называют специалиста в данной области знаний, имеющего положительный опыт решения поставленных перед ним задач.

Несмотря на очевидное достоинство (на данном этапе развития ни одна искусственно созданная вычислительная система не в состоянии смоделировать интеллектуальную работу эксперта, обладающего большим объемом плохо формализуемых и не формализуемых знаний, многолетним опытом работы и интуицией), у экспертного прогнозирования существует ряд явных недостатков.

Для него характерны:

- довольно длительный период составления;
- сбои в работе при временном отсутствии эксперта;
- невозможность быстрой и точной обработки большого объема информации.

В случае применения метода экспертной оценки, следует сказать также о том, что большое влияние на составление прогноза оказывает компетентность эксперта. Под компетентностью эксперта понимают следующее: это наличие у эксперта набора определенных знаний, опыта в данной конкретной области, позволяющих ему выносить верные суждения о возникшей проблеме и предпринимать эффективные меры для ее разрешения.

Почасовое потребление электрической энергии объектами, которые принято называть группами точек поставки, имеет весьма сложный характер. Временные графики этого потребления не всегда обладают определяемой на глаз регулярностью. Для таких систем статистически допустимые возмущения несут вес, сравнимый с регулярными составляющими потребления.

В процессе прогнозирования можно применять разные алгоритмы и методики, но этот выбор никак не влияет на трудозатраты, необходимые для получения реального результата. Прогнозирование потоков электроэнергии - непростая задача, в основе которой лежит комплекс действий по накоплению

достаточной статистики и анализу данных. В отличие от научной лаборатории, где опыты проводятся без изменения самого объекта, в случае с электропотреблением объект способен меняться, притом кардинально, о чем исследующие его сотрудники зачастую узнают не сразу. Изменения состава или конфигурации объекта исследования порой приводят к аннулированию уже полученных результатов.

На практике столь сложная экспериментально-аналитическая задача стоит не перед профессиональной научной организацией, а перед обычными инженерами. В общем случае, проведение квалифицированного прогнозирования специалистами на местах крайне затруднительно, к тому же решение подобных вопросов не входит в их основные должностные обязанности.

Понимая всю сложность формирования прогнозов, руководители объектов включают в технические задания на разработку соответствующих систем требования полной автоматизации процедуры прогнозирования. Это позволяет им снять с себя ответственность за качество прогноза.

Первое, что необходимо выбрать – это метод прогнозирования, т.е. способ получения прогнозных значений электропотребления. Здесь стоит отметить, что выбор метода прогнозирования является лишь начальным этапом в создании методики прогнозирования. Под методикой прогнозирования в общем случае понимается набор решений и правил, позволяющих составлять прогноз, отвечающий требованиям поставленной задачи.

К настоящему моменту разработано несколько основных методик прогнозирования, которые в своей основе имеют методы статистического анализа. Перечислим основные из них:

- моделирование;
- экстраполяция;
- эвристический метод составления прогноза;
- регрессионные методы прогнозирования.

Для прогнозирования электропотребления очень часто применяется так называемый эвристический метод.

Под методом эвристического прогнозирования в общем случае понимается опрос высококвалифицированных специалистов в данной области знаний, позволяющий составить целостную картину возможных решений и развития конкретной ситуации. В данном случае выдвигаемые экспертами суждения основываются на их знаниях, опыте, и, что немало важно, профессиональной интуиции. В этой связи справедливо отметить, что все вышеперечисленные факторы делают эвристическое прогнозирование очень субъективным, т.е. практически полностью зависимым от мнения конкретного человека. Этот фактор становится очень весомым особенно в том случае, если отсутствует возможность опроса нескольких экспертов и решение задачи прогнозирования возложено на одного человека.

Эвристическое прогнозирование можно назвать своего рода мозговым штурмом, т.к. в основе этого метода лежит коллективное принятие решения, путем высказывания каждым экспертом своего видения ситуации и ее развития. Затем происходит выработка оптимального решения, которое наиболее точно и адекватно могло бы спрогнозировать развитие сложившейся ситуации. В нашем случае, это график электропотребления. Стоит отметить, что отличие эвристического метода прогнозирования от мозгового штурма заключается в том, что в данном методе для составления суждений применяются более формализованные средства, такие как таблицы, опросные листы и т.д. В случае мозгового штурма все эти формальности отсутствуют.

Данный метод является наиболее ранним и наиболее простым, с технической точки зрения и нашел широкое применение во всех областях жизни человека.

Одной из особенностей данного метода можно назвать то, что сам эксперт не отдает себе отчета в том, на основании каких конкретных данных он принимает то или иное решение, однако, это не делает эвристическое

прогнозирование примитивным, а напротив, позволяет достигать высоких результатов в части точности прогноза. Если рассматривать сферу электроэнергетики, то здесь это наиболее заметно, т.к. не смотря на высокий уровень развития техники и информационных технологий, прогнозы составленные экспертами в большинстве случаев превосходят те, что составлены на базе специализированных компьютерных программ или искусственного интеллекта.

Высокие результаты эвристического метода прогнозирования можно объяснить тем, что постоянно совершенствуются методы структурирования оценок экспертов, а также высокой квалификацией самих экспертов, что является решающим фактором в определении точности прогноза.

К основным методам экспертной оценки можно отнести:

- матричный метод;
- метод коллективной генерации идей;
- метод комиссии.

Основной целью эвристического прогнозирования является создание наиболее объективной модели развития конкретной ситуации, в нашем случае это потребление электроэнергии в краткосрочной перспективе, на основании мнений квалифицированных специалистов в данной области. Для достижения поставленной цели необходима также и правильная обработка экспертных оценок, которая позволяла бы отсеивать случайные и маловажные факторы, а концентрироваться на наиболее весомых обстоятельствах, влияющих наибольшим образом на конечный результат прогноза.

В основе эвристического прогнозирования лежат следующие допущения:

- эксперт способен на основании своего профессионального опыта, интуиции и знаний в данной конкретной области составить свое мнение о будущем и выразить его в приемлемой для понимания и дальнейшей работы форме;

- схожесть процесса прогнозирования методом экспертной оценки и поиском решения научной проблемы в том смысле, что формы представления решений и в том, и в другом случаях одинаковы и носят характер моделирования тенденций развития в ближайшей временной перспективе;

- способности составления прогнозных моделей на основании объективного понимания проблематики данной области и возможных путей решения проблемы, а также способность выбора наилучшей из представленных моделей.

Все эти допущения могут быть реализованы в методе экспертной оценки только благодаря тому, что существуют методики обработки этих оценок, позволяющие на выходе получить адекватную поставленной задаче модель прогнозирования.

Но эвристические методы необъективны и применяются только тогда, когда существуют эксперты, которые хорошо знакомы ситуацией подвергаемой прогнозу. К тому же, при прогнозировании характеристик сложных технических объектов, методы эвристического прогнозирования становятся очень сложными и трудоемкими.

Этот факт потребовал разработки математических методов прогнозирования. Главные достоинства таких методов заключаются в объективности полученной информации и в возможности автоматизации процесса вычисления прогноза с использованием электронных вычислительных машин. К их числу можно отнести методы прогнозирования при помощи экстраполяции.

1.2 Статистические методы прогнозирования

Экстраполяция является одним из наиболее популярных и эффективных методов прогнозирования. Суть данного метода заключается в том, что результирующий график функции представляется в виде суммы

двух слагаемых - периодической и случайной. Математическое описание данного определения представлено в формуле (1):

$$y(x) = f(\bar{a}, x) + \eta(x); \quad (1)$$

Регулярная составляющая $f(a, x)$ представляет собой в общем случае гладкую функцию от аргумента (времени), которую описывает вектор a , сохраняющий свое значение на всем протяжении составления прогноза. Эту составляющую также можно назвать линией тренда. Это означает, что подобная прямая лишена случайных составляющих и содержит в себе только значения, описывающие сущность процесса и учитывающие только основные факторы, повлиявшие на ход его развития. Однако стоит отметить, что существует множество процессов, в которых порой очень затруднительно разделить случайную составляющую от тренда, поскольку сложно предсказать величину влияния того или иного фактора на развитие всего процесса. К таким процессам можно отнести и электропотребление крупного предприятия, каким и является энергосбытовая компания. Необходимо также сказать, что разделение тренда и случайной составляющей должно происходить исходя из конкретных целей и требуемых показателей точности прогнозирования.

Случайная составляющая $\eta(x)$ – это независимый случайный процесс, чье математическое ожидание равно нулю. Эту составляющую необходимо учитывать для увеличения точности прогнозирования.

Сущностью метода экстраполяции прогнозируемой величины является нахождение такой линии тренда, которая бы наилучшим образом описывала бы происходящий процесс и наиболее полно учитывала бы факторы, влияющие на его ход. Необходимо заметить, что экстраполяционный метод прогнозирования схож с составлением регрессионной модели, но в то же время имеет и свои специфические черты и способы составления прогноза,

вследствие чего его можно отнести к самостоятельным методам прогнозирования.

К особенностям данного метода прогнозирования можно отнести то, что он требует предварительной подготовки числового ряда, т.е. приведения его к виду, удобному для составления прогноза. В этом же ряду необходимо назвать и то, что предварительно происходит анализ физики и логики прогнозируемого процесса. Это влияет на выбор функции экстраполяции и параметры, определяющие границы ее изменения.

Предварительная подготовка и физическая обработка числового ряда направлена в первую очередь на то, чтобы снизить влияние случайной составляющей в исходном числовом ряду и как можно более приблизить его к линии тренда. Второй задачей является снижение трудоемкости математического описания тренда. Для этого чаще всего применяются процедуры сглаживания и выравнивания статистического ряда.

Процедура сглаживания имеет своей целью уменьшение случайных отклонений значений временного ряда от линии тренда, которая в общем случае наилучшим образом описывает основное течение процесса. Самым распространенным методом сглаживания кривой является осреднение значений соседних точек. Эта процедура происходит на всем протяжении числового ряда и поэтому ее зачастую называют скользящая средняя. Для более простых случаев применяют линейную сглаживающую функцию, которая усредняет значения временного ряда, находящиеся в интервале двух ближайших значений. Существуют также нелинейные функции усреднения, но принцип их действия сложнее и они усредняют значения временного ряда на интервале, включающем в себя гораздо большее количество точек.

Принцип действия сглаживания заключается в том, что с помощью многочленов находятся наименьшие квадраты из группы опытных точек. Наилучшее сглаживание получается для точек, находящихся в середине группы, поэтому количество точек в группе обычно выбирается нечетным.

Процедура сглаживания, даже в ее линейном варианте, является очень эффективным способом отделения линии тренда от случайных помех и шума в данных. Имеется возможность проводить процедуру сглаживания несколько раз. Это особенно эффективно для числовых рядов, в которых имеются отклонения от линии тренда со значительной амплитудой. Количество последовательных процедур сглаживания может быть различным. Оно определяется видом числового ряда, целью, которую необходимо достичь при помощи сглаживания, а также величиной приближения к линии исходных данных. Стоит также отметить, что эффективность процедуры сглаживания быстро падает от раза к разу, поэтому рекомендуется проводить ее от одного, до трех раз, в зависимости от достижения поставленных целей сглаживания.

Если говорить о линейном сглаживании, то необходимо сказать, что это считается грубой процедурой, которая способна выявить лишь приблизительный вид линии тренда данных. Для того чтобы более точно определить форму сглаженной кривой, необходимо применять операцию нелинейного сглаживания или как ее еще называют, нелинейные скользящие средние. При такой операции ординатам точек, входящих в группу скользящих, присваивают различные веса в зависимости от их удаленности от середины интервала сглаживания.

Отличие процедуры выравнивания от процедуры сглаживания состоит в том, что при сглаживании выявляется основная линия тренда данных и отсеиваются помехи и случайные отклонения в них, в то время как при выравнивании значения исходных данных остаются такими же, но форма их представления изменяется на более удобную для решения конкретных задач.

Самыми распространенными процедурами выравнивания являются процедуры логарифмирования и замены переменных.

В том случае, если исходная формула имеет три переменные или же известно, что функция трехпараметрическая, то после проведения

определенных математических преобразований зачастую удается избавиться от одного из параметров, а два других привести к формулам выравнивания.

Выравнивание необходимо рассматривать не только как представление исходных данных в наиболее приемлемой форме, но и как способ нахождения приближенных параметров функции, аппроксимирующей данный исходный числовой ряд. Данный метод очень часто используют именно для определения этих параметров при прогнозировании. Вообще возможность применения данного метода для определения параметров аппроксимирующей функции исходных данных определяется характером самих исходных данных, знаний эксперта и его способностью описать заданный процесс с использованием такой функции.

Зачастую вид функции, описывающий прогнозируемый процесс, неизвестен. В этом случае следует рассматривать операцию выравнивания как подготовительную часть, позволяющую сделать выбор, в каком виде наиболее удобно было бы представить ее для решения задачи прогнозирования.

К методам выравнивания также можно отнести и исследование эмпирического числового ряда, для выявления некоторых особенностей функции, аппроксимирующей его. Далеко не всегда указанные преобразования могут привести к получению линейных форм функции. Для задач прогнозирования результаты таких преобразований очень важны, поскольку облегчают выбор аппроксимирующей функции. Для простейших случаев обычно используют следующие типы дифференциальных функций:

- первая производная;
- логарифмическая производная;
- эластичность функции.

Далее более подробно остановимся на рассмотрении статистических методов прогнозирования.

Статистические методы прогнозирования — это область научного знания, основной целью которой является изучение, разработка и

использование на практике современных статистических методов прогнозирования на основе имеющихся данных. Также к задачам этой области научного знания можно отнести развитие и практическое применения методов экспертного прогнозирования. Она изучает также прогнозирование в условиях риска и совместное применение методов экспертной оценки и математико-статистических методов прогноза.

Базируются статистические методы прогнозирования на теории принятия решений и прикладной математической статистике.

Необходимым этапом прогнозирования является оценка точности прогноза. Зачастую это делают при помощи доверительных интервалов. Также для этих целей строят наилучший прогноз по методу максимального правдоподобия, что в свою очередь можно отнести к вероятностно-статистическим моделям восстановления зависимостей. Существуют также параметрические и непараметрические методы оценки прогнозирования. К параметрическим методам можно отнести метод нормальных ошибок, а к непараметрическим методы на основе центральной предельной теоремы теории вероятности. Также нашли свое применение и эвристические приемы, о которых мы говорили ранее, но которые не основаны на вероятностно-статистической теории. Это метод скользящих средних и метод экспоненциального сглаживания.

В настоящий момент можно говорить о том, что основным статистическим аппаратом прогнозирования является многомерная регрессия, которая использует непараметрические оценки плотности распределения. Стоит сказать о том, что можно не использовать предположение о нормальности погрешности измерений и отклонений от линии регрессии, однако для того чтобы это сделать, необходимо взять в качестве основы иной математический аппарат, такой как многомерная центральная предельная теорема теории вероятности, операции линеаризации и наследования сходимости. Это позволяет строить доверительные оценки прогнозирования, а также производить точечное и

интервальное оценивание параметров и проверять значимость их отличия от 0.

Очень важной проблемой при прогнозировании временных рядов является проблема адекватности модели, а также проблема отбора факторов, влияющих на результаты прогнозирования. Зачастую список таких факторов весьма широк и первоочередной задачей является его сокращение, насколько это возможно. Поэтому современные исследования в области прогнозирования направлены именно на то, чтобы выявить именно те признаки, которые оказывают наибольшее информативное влияние на построение и точность прогнозной модели. В данный момент исследователи ищут наилучшие способы решения данной проблемы. Можно говорить о том, что наиболее перспективными непараметрическими методами являются методы оценивания плотности вероятности, а также применение их для восстановления регрессионной зависимости. Наилучших результатов в решении этой проблемы удалось достичь благодаря применению подходов статистики нечисловых данных.

Если говорить о других перспективных направлениях исследования в области статистических методов прогнозирования, то необходимо отметить, что на данный момент к таким направлениям можно отнести авторегрессионный анализ, а также системы эконометрических уравнений основанных как на параметрических, так и непараметрических подходах.

Необходимо сказать о том, что для успешного построения прогнозных моделей в настоящий момент невозможно обойтись без использования современных информационных технологий, позволяющих применять компьютерные статистические методы прогнозирования. Тенденция на данный момент такова, что компьютерные технологии позволяют объединять различные методы прогнозирования на одном рабочем месте прогнозиста, что в свою очередь дает ему свободу выбора наиболее эффективной модели, а также сокращает время составления прогноза.

Составление прогноза качественных признаков, имеющих нечисловую природу, основано на статистическом анализе нечисловых данных. Перспективными для такого прогнозирования представляется прогноз, составленный на основе регрессионного анализа интервальных данных, включающий в себя расчет и определение рационального объема выборки, а также регрессионный анализ нечетких данных. Необходимо сказать о том, что полезным является применение дисперсионного анализа и дискриминантного анализа для реализации на программном уровне статистических методов прогнозирования. Данные виды анализа являются частными случаями регрессионного анализа нечисловых данных.

1.3 Методы прогнозирования, основанные на использовании искусственных нейронных сетей

В настоящий момент большой интерес выражается к методам, которые основываются на нелинейных моделях. Большинство таких методов относится к классу технологий искусственного интеллекта.

В последнее время нейронные сети все больше используют и совершенствуются во всех сферах человеческой деятельности, недаром популярность нейронных сетей набирает все большие обороты.

Для того чтобы понять принцип действия искусственных нейронов, необходимо сначала разобраться из чего состоят и каким образом действуют нейроны человеческого мозга. Необходимо начать с понятия развития нейронов и здесь никак нельзя обойти понятие пластичности (plasticity). Это такая способность мозга, которая позволяет ему выстраивать нервную систему в соответствии с окружающими условиями. Именно эта способность играет самую важную роль при обработке информации нейронами человеческого мозга. В искусственных нейронных сетях работа нейронов происходит аналогичным образом. Искусственная нейронная сеть (neural network) в общем случае представляет собой аппарат, позволяющий моделировать различные способы обработки информации человеческим

мозгом. Подобные сети обычно реализуются на уровне электронных компонентов или же на базе моделей, создаваемых в специализированных компьютерных программах.

В данной главе магистерской диссертации производится описание и сравнение принципов работы экспертных и статистических методов прогнозирования, с методами прогнозирования, базирующимися на работе искусственных нейронных сетей, которые осуществляют составление прогноза в ходе процесса обучения (learning). Искусственная нейронная сеть, также как и человеческий мозг, способна добиваться высоких результатов производительности не только и не столько благодаря большому количеству нейронов в сети, а сколько благодаря огромному количеству связей между всеми нейронами сети. Учитывая все вышесказанное, можно дать следующее определение искусственной нейронной сети:

Искусственная нейронная сеть — это большой параллельный распределительный процессор, состоящий из простейших частиц (моделей нейронов) обработки информации, который способен накапливать знания, полученные в ходе обучения, и применять их для решения задач, схожих по принципу с обучающими, но не входящих в их число. Можно выделить следующие свойства, которые присущи и человеческому мозгу и ИНС:

- процесс обучения заключается в получении знаний из окружающей среды и дальнейшей их обработке при помощи нейронов;
- накопление знаний происходит при помощи создания особых связей между нейронами, называемыми синаптическими весами.

На данный момент применение искусственных нейронных сетей достигает все больших масштабов, так как зачастую использование обычных алгоритмических или статистических методов оказывается неэффективным или вовсе невозможным. К областям, в которых широко применяются искусственные нейронные сети можно отнести распознавание текстов, контекстная реклама в интернете, прогнозирование погодных условий, прогнозирование котировок валют и т.д.

Применительно к теме данной диссертации необходимо сказать, что краткосрочное прогнозирование электропотребления является перспективной областью использования искусственных нейронных сетей. Все больше производителей программного обеспечения выходят на рынок с программными продуктами, способными составлять подобные прогнозы.

С самого начала компьютерной эры существовали два принципиально разных подхода к обработке информации, а именно: последовательная обработка символов и параллельное распознавание образов. Различие символов и образов состоит лишь в размерности, однако, принципы их обработки компьютером остаются теми же. Стоит заметить, что размеры образов могут быть на много порядков больше, чем размеры символов. На первый взгляд может показаться, что различия в размерности приводят лишь, к тому, что увеличивается время обработки такой информации компьютером, однако, на деле оказывается, что с возрастанием размерности информационного блока возрастает и сложность его обработки компьютером. Причем это возрастание сложности обработки происходит нелинейно.

Простота обработки коротких символов происходит из-за того, что возможно создать такой процессор, который бы предсказуемым образом обрабатывал все входящие символы, которые можно назвать данными или сигналами, и производил над ними все возможные операции. В случае же с образами сделать такой процессор не представляется возможным, поскольку различных образов может быть огромное количество, и для того чтобы их все запомнить потребовался бы процессор огромной мощности. Отсюда естественно следует вывод о том, что необходимо вместо увеличения мощности процессора, изменить способ обработки информации этим процессором. Следовательно, процессор, предназначенный для обработки образов, должен обладать лишь минимальным набором примеров образов, который бы в то же время позволял ему самостоятельно разрабатывать стратегию поведения и принятия решений. И в то же время эти решения

должны быть аналогичными тем, которые были получены им в ходе обучения, а также должны быть адекватными поставленной задаче. Из всего вышесказанного можно сделать вывод о том, что различие между последовательными и параллельными вычислениями состоит в принципиально разном подходе к обработке входящей информации.

Компьютеры, реализованные на принципах последовательной обработки символов ограниченной длины, работают на архитектуре фон Неймана, что является традиционным подходом для такого типа машин. В то время как нейрокомпьютеры работают по принципам параллельного вычисления и распознавания образов, что делает их работу схожей с работой мозга человека. Для того чтобы еще раз показать различия в способах обработки информации компьютерами с последовательным принципом работы и человеческим мозгом, работающим по принципу распознавания образов, можно сказать, что такие компьютеры обладают большой мощностью при нахождении численного решения задачи, и затрачивают на это доли секунды в отличие от человека, которому для этого понадобится значительное количество времени. Однако, например, узнать лицо знакомого для человека не составляет труда и это происходит практически мгновенно, в то время как для компьютера с последовательным принципом вычислений это становится практически непосильной задачей.

Из всего вышесказанного следует, что задача нейрокомпьютеров это обработка образов. У таких компьютеров нет разделения на активный процессор и пассивную память и все вычисления и процесс обучения распределены по всем нейронам сети. Этим они схожи по строению с мозгом человека. Эти факторы позволяют нейрокомпьютерам добиваться очень высокой производительности, которая может превышать производительность компьютеров с последовательными вычислениями на много порядков.

Преимущества нейросетевой архитектуры над параллельными вычислениями:

- способность к параллельной обработке информации;

- единый принцип обучения для всех единиц обработки информации (нейронов);

- при выходе из строя какой-либо части нейронов, компьютер продолжает функционировать, что значительно повышает его надежность;

- такие компьютеры способны решать нестандартные задачи.

Различие в развитии человека и компьютеров можно описать следующим образом: человек, рождаясь, учится в начале распознавать образы, а только потом, многим позже начинает учиться мыслить логически. В развитии компьютерной техники происходит все наоборот. С самого начала в компьютеры закладывался принцип последовательных вычислений, и только потом, с помощью применения специальных программ, позволяющих эмулировать нейросетевую архитектуру на процессорах с последовательным принципом вычислений, машины становятся способными к распознаванию образов. Также существуют и аппаратные методы построения архитектуры параллельных вычислений, но они значительно сложнее в реализации, нежели создание специализированного программного обеспечения.

Появление вычислительной техники совершенного нового класса (нейрокомпьютеров) обусловлено с одной стороны ускоренным развитием элементной базы, а с другой – необходимостью ускорения решения конкретных математических задач.

Развитие нейрокомпьютеров было обусловлено развитием пороговой логики. Название же свое они получили благодаря условному названию порогового элемента с фиксированными или настраиваемыми весами, реализующего простейшую передаточную функцию по принципу нейронов человеческого мозга.

Значительный скачок в развитии нейронных ЭВМ произошел в 1980-х годах, благодаря появлению возможности эмуляции нескольких элементов-нейронов и всего множества связей между ними в одном или нескольких

кристаллах, что ранее было невозможно. Это позволили сделать электронные и оптические методы реализации вычислительных систем.

Основным принципом работы нейрокомпьютеров является аналогово-цифровое построение ЭВМ. При таком построении ЭВМ, быстрая аналоговая часть выполняет необходимые вычисления в пороговом элементе. В свою очередь, алгоритмы настройки весов нейронных сетей реализуются в специализированных цифровых системах.

Логической основой вычислительной системы является группа операций, составляющая основную часть, и реализуется она на основных элементах операционного устройства. Если рассматривать нейрокомпьютеры, логический базис вычислительной системы делает необходимым наложение сложной системы программирования на него. В самом простом и чаще всего встречающемся случае, базисом вычислительной системы становится базис $\{\sum \bar{a}x, sign\}$, который максимально соответствует базису решаемых задач. При таком подходе, решение задач с логическим базисом, описанным выше будет происходить в системе с таким же базисом. Это позволяет избежать искусственных сдвигов. Данный факт должен обеспечивать максимальную производительность при решении задачи.

Также одной из особенностей нейронных ЭВМ является то, что ее структура может быть рассчитана аналитически, в то время как структура обычных машин получается эмпирическим путем, исходя из субъективных представлений о решаемой задаче.

Типы решаемых задач, которые возможно прорабатывать на подобных системах, можно разделить на 3 класса:

- формализуемые;
- трудно формализуемые;
- не формализуемые.

К формализуемым задачам относятся задачи, алгоритм решения которых ясно очерчен.

К трудно формализуемым обычно относят такие задачи которые имеют множество путей решения, либо эти решения невозможно оценить с точки зрения их качества или же достижимости этого решения в принципе.

Особый класс задач представляют собой так называемые не формализуемые задачи. Они названы таким образом, потому что в алгоритме их решения присутствуют функции или данные, заданные неявно, т.е. в виде входных сигналов или образов. Перечислим основные виды таких задач:

- кластеризация;
- распознавание образов;
- самообучение;
- прогнозирование.

Подобие искусственных нейронных сетей и человеческого мозга не вызывает сомнений, поскольку оба этих процессора имеют похожую структуру, состоящую из многих мельчайших процессоров, функционирующих параллельно и имеющих множество связей друг с другом. Не удивительно, что при такой схожей структуре этих сетей, они демонстрируют и схожесть своих способностей, таких как обучение на основе опыта решения обучающих примеров, обобщение, отделение значимых данных от посторонних шумов и помех во всем потоке информации. И это не смотря на то, что при создании искусственных нейронных сетей делается ряд значительных допущений.

Нейронные сети также обладают способностью перестраиваться в случае изменения внешних условий окружающей среды. Происходит это следующим образом: в случае изменения входного сигнала, ИНС само настраивается и обучается таким образом, чтобы выходной сигнал обеспечивал правильную реакцию на изменение окружающей среды. Правильно обученная нейронная сеть становится устойчивой к незначительным отклонениям входного сигнала от нормы. Это позволяет ей продолжать видеть нужный «образ» не смотря на зашумленность потока входных данных.

Чтобы построить нейросетевую модель, способную адекватно и с заданной точностью решать поставленную перед ней задачу, необходимо:

1. Описать объект, выступающий в качестве входного сигнала нейронной сети. Это могут быть исходные значения переменных, начальные значения определяемых величин. В нашем случае в роли входного сигнала будут выступать данные об электропотреблении за предыдущий аналогичный период.

2. Определить объект, который будет выступать в качестве выходного сигнала. Это могут быть искомые значения переменных. В случае с краткосрочным прогнозированием электропотребления, это будет график потребления на сутки вперед с почасовой разбивкой.

3. Определить так называемый целевой выходной сигнал нейросети, т.е. такую величину электропотребления, к которой нейронная сеть должна стараться приблизить свой прогноз.

4. Определить архитектуру искусственной нейронной сети:

- а) количество слоев;
- б) способ связи между слоями;
- в) весовые коэффициенты.

5. Определить функцию, с помощью которой нейронная сеть «исправляла» бы отклонения реального выходного сигнала от желаемого.

6. Определить критерии оценки качества работы сети и способы ее оптимизации.

7. Определить значения синаптических весов для нейронов сети. Для этого существуют различные способы, такие как:

- а) аналитический способ, при непосредственной постановке задачи;
- б) при помощи численных методов;
- в) воспользовавшись процедурой настройки коэффициентов искусственной нейронной сети.

Решение задачи с помощью искусственной нейронной сети предполагает работу этой сети в определенном режиме и обработку конкретных числовых данных.

Процедура решения задачи включает в себя:

1. Разработку структуры нейронной сети, которая отвечала бы поставленной задаче;
2. Вычисление весовых коэффициентов нейронной сети или выбор их из памяти, если они были вычислены ранее. Последнее свойство будет в дальнейшем применено в этой работе при прогнозировании электропотребления;
3. При необходимости, генерацию начальных или исходных параметров;
4. Загрузку всех численных значений в ИНС и запуск расчета;
5. Работа сети в соответствии с режимом обучения:
 - а) за фиксированное число шагов;
 - б) за переменное число шагов, в ходе которых происходит минимизация ошибки сети, т.е. таким образом сеть настраивает весовые коэффициенты таким образом, чтобы отличие выходного сигнала от желаемого было минимальным.
6. Генерация решения.

В случае многократного использования сети, шаги 1 и 2 повторять не обязательно.

Наиболее распространенным подходом к проектированию искусственных нейронных сетей является анализ и выбор структуры сети из уже имеющихся шаблонов (сети Хопфилда, Гросберга, Кохонена), исходя из класса поставленной задачи. По большому счету, в каждой конкретной задаче структура нейронной сети изменяется незначительно, по сравнению с типовой структурой. Возможно лишь изменение синаптических весов, способов обучения, способов представления входных и выходных данных.

Исходной точкой справедливо можно считать формулировку задачи. Исходя из нее, выбирается структура сети, адекватная этой задаче. Выбор делается из уже существующих структур, разработанных под конкретный тип задач. Однако, нельзя говорить, что типовая структура будет применима к конкретной задаче, поэтому обязательно требуется настройка сети, для получения адекватного и точного решения.

1.4 Разомкнутые искусственные нейронные сети

Обычно при анализе разомкнутых искусственных нейронных сетей используется метод вероятностного анализа многомерных нелинейных систем. Использование этого метода обуславливается тем, что при его использовании результаты не зависят от сложности структуры и разновидности нейронной сети. Результаты такого анализа могут зависеть только от характеристик пространства признаков и пространства решений. Эта особенность в дальнейшем используется для формирования функционала вторичной оптимизации нейронной сети, а также для построения замкнутых нейронных сетей.

Функционал вторичной оптимизации нейронной сети - это функционал, который выражен через параметры распределений текущих сигналов и минимизируем в многослойных системах.

На этом этапе рассматриваются две проблемы синтеза.

Первая проблема заключается в том, что необходимо исследовать соответствие функционала вторичной оптимизации и критериев первичной оптимизации. И в качестве примера здесь можно рассматривать известные адаптивные структуры, такие как матрица Штайнбуха или же трехслойный персептрон Розенблатта.

Основным недостатком такого подхода является то, что не выясняется, каким именно критериям первичной оптимизации соответствует функционал вторичной оптимизации. Это приводит к тому, что некоторые системы

становятся неработоспособными вследствие много модальности поступающих данных.

Второй и основной проблемой на данном этапе синтеза является проблема формирования функционала вторичной оптимизации таким образом, чтобы он соответствовал конкретному критерию первичной оптимизации. Под соответствием здесь понимается совпадение параметров искусственной нейронной сети при условии минимума функционалов вторичной и первичной оптимизации.

На сегодняшний день разработано большое количество структурных форм нейронных сетей и каждая из них ориентирована на решение конкретной задачи. Для наглядности рассмотрим многослойную полно связанную нейронную сеть прямого распространения входного сигнала (рисунок 1), основной функцией которой является распознавание и классификация образов. Термин полно связанная обозначает, что все нейроны слоя i связаны со всеми нейронами слоя $i+1$, а в случае первого слоя – со всеми входами нейронной сети. Прямое распространение данных означает, что в сети отсутствуют обратные петли, по которым бы данные, пройдя, например, через первый слой сети, вновь поступали бы на соответствующие входы.

Особенностью такой сети является то, что она имеет ограниченные возможности реализации разнообразных конфигураций разделяющей поверхности. Это происходит вследствие того, что нейрон не способен реализовать булевскую функцию от $N > 2$ переменных. Если число нейронов во втором (скрытом) слое не ограничено, то для создания кусочно-линейной поверхности достаточно иметь 3 слоя искусственной нейронной сети.

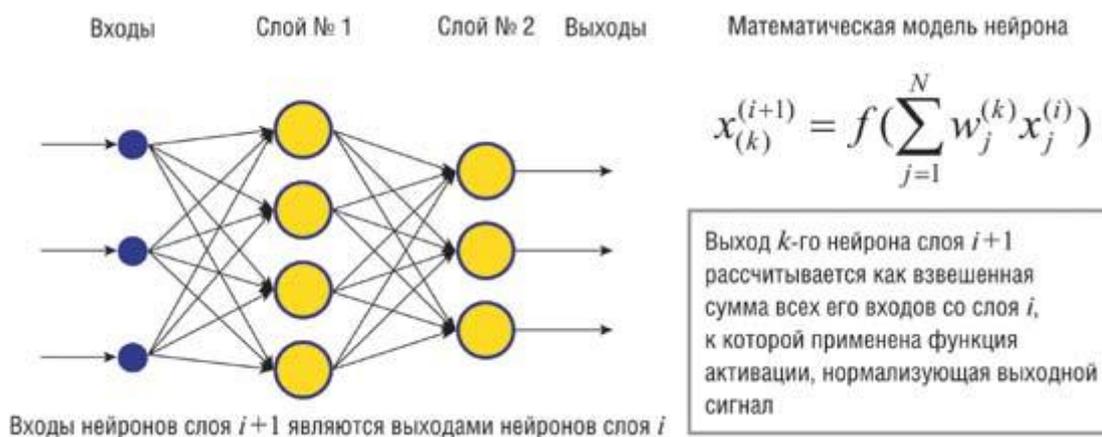


Рисунок 1 - Пример многослойной полно связанной нейронной сети прямого распространения сигнала.

В общем случае, постановкой задачи обучения является процесс, при котором функция, заданная учителем, аппроксимируется автоматом с заданными свойствами. В свою очередь, задачу самообучения можно рассматривать как выделение мод функции распределения и распознавания образов на входе. Объединить эти два определения можно, говоря об обучении искусственной нейронной сети учителем, параметры которого также заданы.

В настоящий момент возрастает популярность структурных методов исследования нейронных сетей в сравнении с символическими методами. Так как искусственные нейронные сети обладают свойствами многомерности, многослойности и многоконтурности, то такой метод исследования становится особенно актуальным. Эти свойства присущи в большей степени современным нейронным сетям. В данной главе магистерской диссертации более подробно рассматриваются вопросы структурной оптимизации, т.е. повышение эффективности работы нейронной сети в большей степени за счет применения различных структурных решений, нежели за счет увеличения объема обучающих данных. Вследствие этого, возникает необходимость структурного описания преобразований разомкнутых сетей, а не только в символическом описании этих преобразований. Далее

представлено описание основных типов нейронных сетей в виде уравнений и граф-схем.

1.5 Многослойные нейронные сети с перекрестными связями.

$$y(n) = F[g(n)] = F\left[\sum_{h_1=0}^{H_1} a_{h_1} y_{h_1}(n)\right] = F\left[\sum_{h_1=0}^{H_1} a_{h_1} F[g_{h_1}(n)]\right] = F\left[\sum_{h_1=0}^{H_1} a_{h_1} F\left[\sum_{i=0}^N a_i x_i(n)\right]\right] \quad (1)$$

Многослойные нейронные сети с полно связанной структурой множество признаков j -го слоя ($j= 1, \dots, W$) состоит из выходных сигналов от 1 до $(j - 1)$ и из признаков исходного пространства.

При сравнении сетей с полными перекрестными связями и с полными последовательными связями выясняется, что первые значительно проще, вследствие меньшего числа нейронов. Но это справедливо только при условии, что обе сети реализуют одну и ту же структуру разделяющихся поверхностей.

Для двухслойной искусственной нейронной сети с перекрестными связями справедливо выражение (2):

$$y(n) = F\left[\sum_{h_1=0}^{H_1} a_{h_1} F\left[\sum_{i=0}^N a_{ih_1} x_i(n)\right] + \sum_{j=0}^N a_j x_j(n)\right] \quad (2)$$

Граф-схема данной нейронной сети представлена на рисунке.

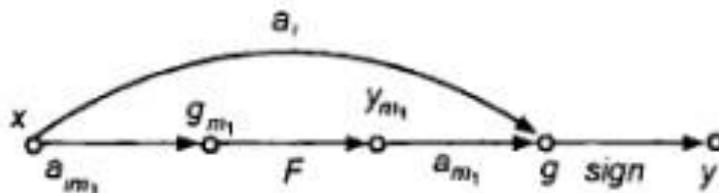


Рисунок 2 - Граф-схема нейрона с обратной связью

В случае двухслойной нейронной сети с обратными связями:

$$y(n) = F[g(n)]; g(n) = \sum_{h_1=0}^{H_1} a_{h_1} y_{h_1}(n) + a_k y(n-1) \quad (3);$$

$$y_{h_1}(n) = F[g_{h_1}(n)]; g_{h_1}(n) = \sum_{i=0}^N a_{ih_1} x_i(n) + a y(n) + a_{h_1} y_{h_1}(n-1) \quad (4).$$

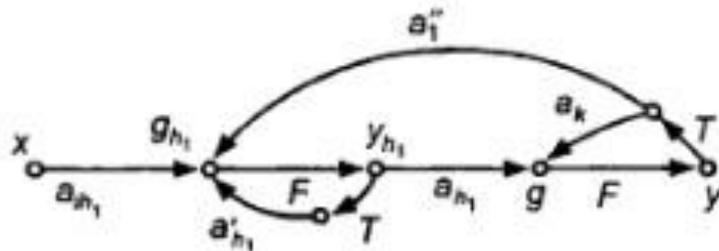


Рисунок 3 - Граф-схема двухслойной нейронной сети с обратными связями

Настоящий уровень развития технологии способен обеспечить эффективную работу нейронной сети, представленной на рисунке 4.



Рисунок 4—Структура многослойной нейронной сети

1.6 Преимущества нейронных сетей

Очевидным является тот факт, что нейронные сети имеют значительные преимущества перед другими методами вычисления, а

именно, это возможность параллельных вычислений, что в разы увеличивает их мощность, а также возможность самообучения, т.е. создания обобщений. Термин обобщение (generalization) означает, что нейронная сеть способна давать адекватное решение задачи, используя при этом данные, которые не встречались в ходе обучения. Все эти свойства делают нейронные сети очень мощным средством вычисления, которые возможно применить для решения крупномасштабных задач. Практическое использование нейронных сетей показывает, что они не способны давать готовые решения при наличии сложной многосоставной задачи. Для того чтобы успешно решать подобные задачи, необходимо их разделять на несколько простых и для каждой из этих задач подбирать соответствующую по структуре сеть.

За счет использования нейронных сетей можно достичь следующих полезных свойств системы:

1. Нелинейность (nonlinearity). Это свойство зависит от того, какими являются сами искусственные нейроны, линейными или нелинейными. В случае если все нейроны сети нелинейны, то и вся сеть считается нелинейной. Причем эта нелинейность распределена по всей сети. Это делает возможным распознавание нелинейных входных сигналов, таких как например человеческая речь.

2. Преобразование входной информации в выходную (input-output mapping). Одной из самых эффективных парадигм обучения является обучение с учителем (supervised learning). Это означает, что при помощи учебных примеров (training sample), изменяются синаптические веса нейронов и сеть перестраивается в соответствии с учебными примерами. Каждый обучающий пример состоит из исходного входного сигнала и желаемого отклика (desired response). Из множества обучающих примеров случайным образом выбирается один и искусственная нейронная сеть выстраивает весовые коэффициенты нейронов таким образом, чтобы отклонение выходного сигнала от желаемого было минимальным.

Этот процесс можно назвать модификацией свободных параметров сети. Однако, стоит сказать о том, что могут использоваться снова и уже имеющиеся весовые коэффициенты, в случае если сеть достигает необходимых результатов точности или же вычисление приходится повторять несколько раз подряд. Обучение нейронной сети или выстраивание необходимым образом весовых коэффициентов нейронов происходит до тех пор, пока отличие выходного сигнала от требуемого не станет минимальным. В ходе обучения таким образом нейронная сеть составляет таблицу, в которой каждому входному сигналу соответствует конкретный выходной сигнал. Это обстоятельство делает обучение нейронной сети похожим на непараметрическое статистическое обучение (nonparametric statistic a inference). Такой тип обучения работает с оценками, которые не соотносятся с конкретными моделями или же с обучением с нуля. Термин «непараметрический» здесь употреблен в том смысле, что изначально не существует предопределенности в связи входных и выходных значений. Для примера рассмотрим задачу классификации образов (pattern classification). В ней требуется соотнести входной сигнал, представляющий физический объект, или событие, с некоторой предопределенной категорией (классом). При непараметрическом подходе к этой задаче требуется «оценить» рамки решения в пространстве входного сигнала на основе набора примеров. При этом не используется никакая вероятностная модель распределения. Аналогичный подход применяется и в парадигме обучения с учителем. Это еще раз подчеркивает параллель между отображением входных сигналов в выходные, осуществляемым нейронной сетью, и непараметрическим статистическим обучением.

3. Адаптивность (adaptivity). Искусственные нейронные сети способны к адаптации во время изменения внешних условий окружающей среды. Даже тот факт, что нейронная сеть обучена в одних условиях окружающей среды, а реальные условия претерпели изменения, не говорит о том, что сеть будет работать ошибочно. Она легко может быть переучена для работы в новых

условиях. Существуют также нестационарные среды, т.е. среды в которых внешние условия окружающей среды изменяются в реальном времени. Также существуют и нейронные сети, чьи весовые коэффициенты изменяются в реальном времени. Типичная архитектура сетей, предназначенных для распознавания образов, задач управления и обработки сигналов, может быть объединена со способностью нейронных сетей адаптации. В этом случае, возможно получить адаптивные сети управления, адаптивные сети распознавания образов и обработки сигналов. Очевидно, что чем выше адаптивные способности нейронной сети, тем устойчивее она может работать в изменяющейся среде. В этой связи, необходимо сказать, что способность к адаптации не всегда работает в положительную сторону. Слишком чувствительная к изменениям внешней среды нейронная сеть способна воспринять помехи в данных как весомую часть входного сигнала. Это может привести к ошибочному формированию выходного сигнала. Идеальным вариантом совмещения способности к адаптации и стабильности может быть сеть, в которой основные параметры, оказывающие наибольшее влияние на выходной сигнал, являются стабильными, но в то же время достаточно пластичными, чтобы иметь возможность реагировать на сильные изменения внешней среды. Эта задача обычно называется дилеммой стабильности-пластичности (*stability-plasticity dilemma*).

4. Очевидность ответа (*evidential response*). Нейронную сеть, предназначенную для распознавания образов можно обучить определять достоверность принимаемого решения. Такую сеть возможно использовать в синтезе с сетями других типов, что позволит повысить их производительность.

5. Контекстная информация (*contextual information*). Нейронная сеть способна накапливать какие-либо знания об окружающей среде только благодаря состоянию ее активации, т.е. параметрам активации пороговых элементов сети. От этих параметров и зависит состояние активации. Поэтому

можно говорить о том, что работа нейронной сети непосредственно связана с контекстной информацией.

6. Отказоустойчивость (fault tolerance). Термин отказоустойчивость означает, что производительность нейронных сетей падает незначительно при неблагоприятных условиях внешней среды. Особенно это относится к сетям, которые на уровне электроники выстроены по принципу нейронных сетей. Это происходит вследствие того, что запомненная информация равномерно распределена по сети и хотя при повреждении одного нейрона или его связей извлечение информации из него затрудняется, снижение производительности всей сети в целом происходит медленно. Поэтому и снижение качества и точности работы нейронной сети происходит медленно. Незначительные повреждения сети вызывают и незначительные отклонения в работе. Это безусловное преимущество вычислений, производящихся по принципам нечеткой логики. Также в искусственные нейронные сети закладываются при проектировании поправки, которые позволяют гарантировать отказоустойчивость.

7. Масштабируемость (VLSI Implementability). Параллелизм вычислений в нейронных сетях обеспечивает ускорение решения задач и масштабируемость нейронных сетей в рамках технологии VLSI (very-large-scale-integrated). Под термином масштабируемость понимается расширение структуры сети для решения более сложных задач, путем добавления в нее нейронов. Это позволяет с помощью технологий VLSI создавать модели сложных процессов, используя иерархические структуры.

8. Единообразие анализа и проектирования (Uniformity of analysis and design). Нейронные сети - это универсальная структура, позволяющая решать задачи из различных областей жизни человека. Причем одно и то же строение сети может быть использовано при решении различных задач. Это свойство проявляется несколькими способами.

- составными частями любой нейронной сети являются нейроны, одинаковые по своей структуре;

- это единообразие позволяет использовать различные алгоритмы обучения на одних и тех же структурах;

- имеется возможность построения модульных сетей, т.е. состоящих из нескольких подсетей, представляющих собой похожие по структуре модули.

9. Аналогия с нейробиологией (Neurobiological analogy). Нейронные сети построены по принципу и подобию с человеческим мозгом. Мозг – это гигантская нейронная сеть, которая обладает высокой быстротой вычислений и большой мощностью. К тому же мозг способен сохранять свою работоспособность при некорректной работе некоторых его областей, например, вследствие получения травм. Все эти свойства вобрали в себя и искусственные нейронные сети. Нейробиологи рассматривают искусственные нейронные сети как средство моделирования физических явлений. С другой стороны, инженеры постоянно пытаются почерпнуть у нейробиологов новые идеи, выходящие за рамки традиционных электросхем. Эти две точки зрения можно продемонстрировать на следующих примерах:

- в работе модели линейных систем вестибуло-окулярного рефлекса сравнивались с моделями рекуррентных нейронных сетей. Вестибуло-окулярный рефлекс, или рефлекс VOR (vestibulo-ocular reflex), является составной частью глазодвигательной системы. Его задачей является обеспечение стабильности визуального образа при поворотах головы за счет вращения глаз. Процесс VOR реализуется премоторными нейронами в вестибулярном центре, которые получают и обрабатывают сигналы поворота головы от вестибулярных сенсорных нейронов и передают результат на моторные нейроны мышц глаз. Механизм VOR хорошо подходит для моделирования, так как входной (поворот головы) и выходной (поворот глаз) сигналы можно точно описать. К тому же это довольно простой рефлекс, а нейрофизические свойства реализующих его нейронов довольно хорошо описаны. Среди трех задействованных в нем типов нейронов премоторные нейроны, входящие в состав вестибулярного центра, являются самыми сложными, а значит, самыми интересными. Ранее механизм VOR

моделировался с помощью сосредоточенной линейной системы и теории управления. Эти модели были пригодны для описания некоторых общих свойств VOR, но не давали четкого представления о свойствах самих нейронов. С появлением нейросетевых моделей ситуация в корне изменилась. Рекуррентные модели VOR (программы, использующие алгоритм рекуррентного обучения в реальном времени) позволили воспроизвести и описать многие статические, динамические, нелинейные и распределенные аспекты обработки сигналов при реализации рефлекса VOR и, в частности, вестибулярный центр;

- сетчатка (retina), более чем какая-то другая часть мозга, выполняет функции взаимосвязи окружающего мира, представленного визуальным рядом или физическим изображением (physical image), проецируемым на матрицу рецепторов, с первым нейронным изображением (neural image). Сетчатка — это матрица микроскопических рецепторов на внешней лицевой стороне глазного яблока. В ее задачи входит преобразование оптического сигнала в нейронное изображение, передаваемое по оптическим нервам в различные центры для анализа. Принимая во внимание синаптическую организацию сетчатки, это — сложная задача. В любой сетчатке преобразование изображения из оптического в нейронное проходит три стадии:

I - снятие фотокопии слоем нейронов-рецепторов;

II - передача сформированного сигнала (реакция на свет) химическими синапсами на слой биполярных клеток (bipolar cell);

III - передача этих сигналов (также с помощью химических синапсов) на выходные нейроны.

На двух последних стадиях (при передаче информации от рецепторов на биполярные рецепторы и от последних — на выходные нейроны) в операции участвуют специальные нейроны с латеральным торможением, в том числе так называемые горизонтальные клетки (horizontal cell). Их задачей является преобразование сигнала между разными синаптическими

слоями. Также существуют центробежные элементы, обеспечивающие передачу сигнала с внутреннего синаптического слоя на внешний. Некоторые исследователи создавали электронные микросхемы, имитирующие структуру сетчатки. Эти электронные чипы назывались нейроморфными контурами (neuromorphic integrated circuit). Нейроморфные сенсоры представляют собой матрицу фоторецепторов, связанных с соответствующими элементами рисунка (пикселями). Они имитируют сетчатку в том смысле, что могут адаптироваться к изменению освещенности, идентифицировать контуры и движение. Нейробиологическая модель, воплощенная в нейроморфных контурах, обеспечила еще одно преимущество: физическое понимание нейробиологических структур может оказать существенное влияние на область электроники и технологию VLSI.

1.7 Сравнение принципов работы человеческого мозга и искусственных нейронных сетей

Нервная система человека является трехступенчатой системой (рисунок 5). Главным центром этой системы является мозг (brain). Его функция состоит в получении информации от всех органов чувств, анализе этой информации и принятии решений к действию, исходя из этой информации. На рисунке 5 стрелки направлены в разную сторону. Направленные слева направо обозначают входную информацию, поступающую через органы чувств, а направленные справа налево, соответствующую реакцию на поступившую информацию. Рецепторы (receptor) преобразуют информацию, поступающую из окружающей среды, в электрические сигналы, передаваемые в нейронную сеть. Функция эффекторов (effectors) заключается в преобразовании электрических импульсов, сгенерированных мозгом, в выходные сигналы.



Рисунок 5 – Блочная диаграмма для нервной системы

Изучение человеческого мозга началось с идеи организации человеческого мозга на основе нейронов. Как правило, реакция нейронов на 5-6 порядков медленнее реакции кремниевых логических элементов. Длительность событий в кремниевых элементах измеряется в наносекундах (10^{-9} с), а в нейронах — в миллисекундах (10^{-3} с). Однако эта относительная медлительность нейронов компенсируется их массой и количеством взаимосвязей между ними. По существующим оценкам, в коре головного мозга насчитывается около 10 миллиардов нейронов и около 60 триллионов синапсов или взаимосвязей между ними. В результате мозг представляет собой чрезвычайно эффективную структуру. В частности, энергетические затраты мозга на выполнение одной операции в секунду составляют около 10^{-16} Дж. В то же время затраты самого экономичного компьютера не опускаются ниже 10^{-6} Дж на операцию в секунду.

Синапсы (synapses) — Это элементарные связи между нейронами, по которым передаются электрические сигналы. Самым распространенным типом синапсов являются химические (chemical synapse), которые работают следующим образом. Предсинаптический процесс формирует передаваемую субстанцию (transmitter substance), которая методом диффузии передается по синаптическим соединениям между нейронами и влияет на постсинаптический процесс. Таким образом, синапс преобразовывает предсинаптический электрический сигнал в химический, а после этого — в постсинаптический электрический. В электротехнической терминологии его можно было бы назвать невзаимным четырехполюсником (nonreciprocal two-port device). В традиционных описаниях нейронной организации синапс представляют простым соединением, которое может передавать возбуждение

(excitation) или торможение (inhibition) (но не то и другое одновременно) между нейронами.

Под пластичностью понимается способность нервной системы адаптироваться к условиям окружающей среды. В человеческом мозге пластичность достигается двумя путями: первый – это создание новых синаптических связей между нейронами, а второй – это совершенствование уже существующих связей. Аксоны (axon) – это линии, по которым передаются электрические сигналы. Дендриты (dendrite) – это зоны приема сигналов, поступающих в нейрон. Эти функциональные единицы представляют собой два типа элементов клетки, которые различаются даже на морфологическом уровне. Аксоны имеют более гладкую поверхность, более тонкие границы и большую длину. Дендриты (свое название они получили из-за сходства с деревом) имеют неровную поверхность с множеством окончаний. Существует огромное множество форм и размеров нейронов, в зависимости от того, в какой части мозга они находятся. На рисунке показана пирамидальная клетка (pyramidal cell) — самый распространенный тип нейронов коры головного мозга. Как и все нейроны, пирамидальные клетки получают сигналы от шупалец дендритов (фрагмент дендрита показан на рисунке 7). Пирамидальная клетка может получать более 10000 синаптических сигналов и проектировать их на тысячи других клеток.

Выходные сигналы большинства нейронов преобразуются в последовательность коротких электрических импульсов. Эти импульсы, называемые потенциалами действия (action potential) или выбросами (spike), берут свое начало в теле нейрона и передаются через другие нейроны с постоянной скоростью и амплитудой. Причина использования потенциала действия для взаимодействия нейронов лежит в самой физической природе аксона. Аксон нейрона имеет большую длину и маленькую толщину, что выражается в его большом электрическом сопротивлении и емкости. Обе эти характеристики распределены по аксону. Таким образом, его можно

смоделировать как линию электропередачи с использованием уравнения кабеля (cable equation). Анализ этого уравнения показывает, что подаваемое на один конец аксона напряжение экспоненциально уменьшается с расстоянием, достигая на другом его конце малых значений. Потенциалы действия позволяют обойти эту проблему.

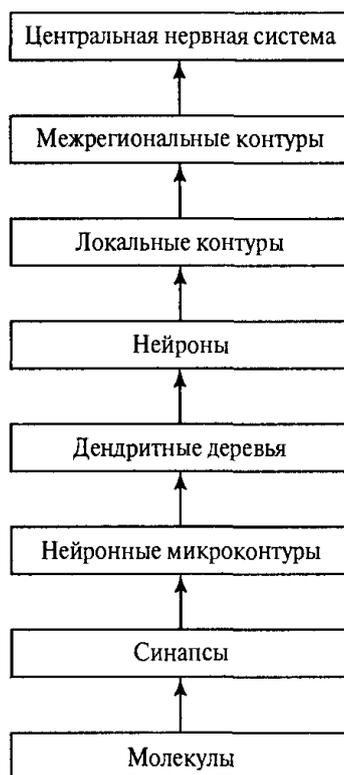


Рисунок 6 – Структурная организация уровней мозга

В человеческом мозге существуют крупно- и мелкомасштабные анатомические структуры. Эти верхний и нижний уровни отвечают за выполнение разных функций. На рисунке показана иерархия уровней организации мозга, составленная на основе анализа отдельных областей мозга.

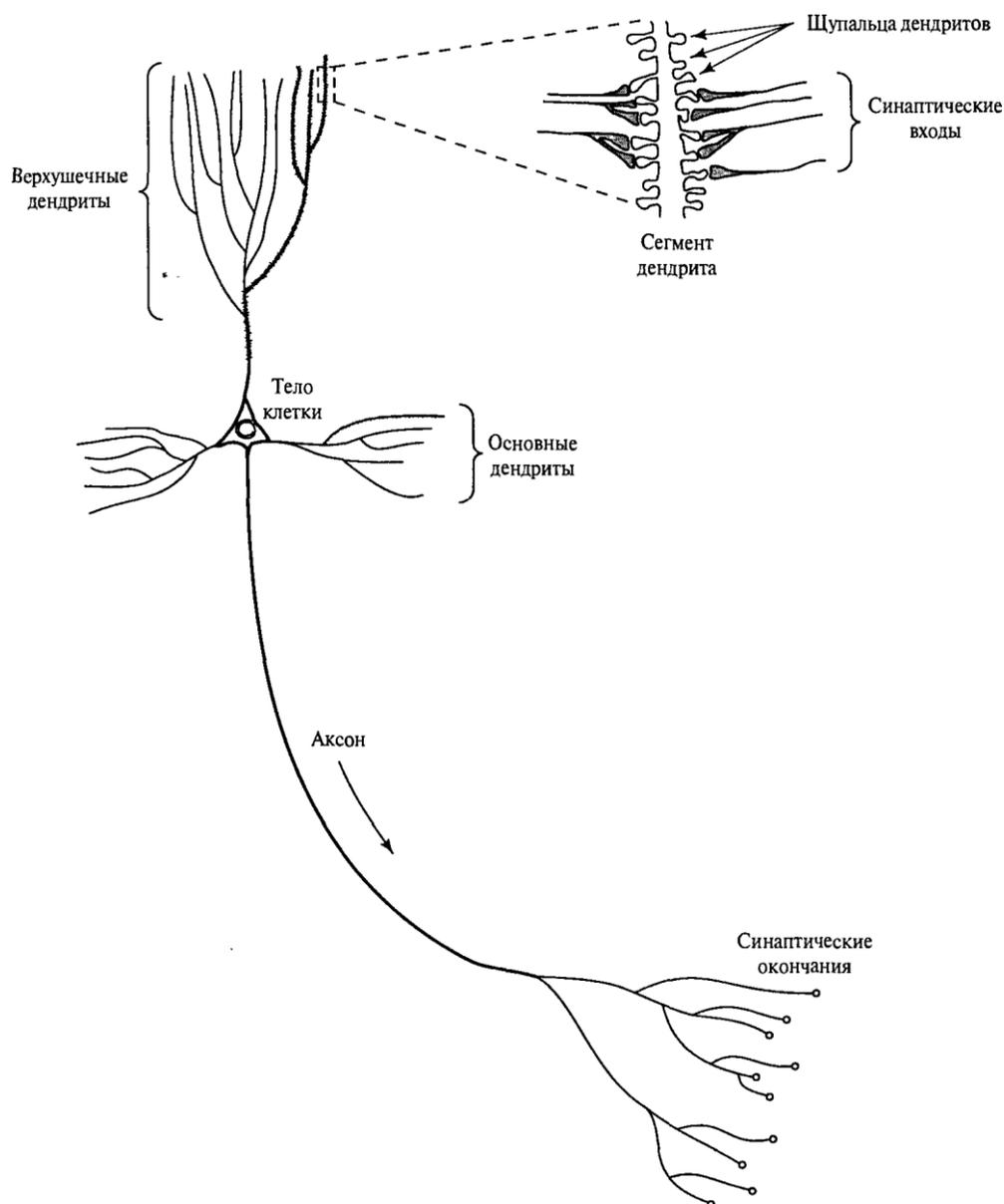


Рисунок 7 – Пирамидальная клетка

Синапсы (synapse) представляют собой самый нижний уровень — уровень молекул и ионов. На следующих уровнях мы имеем дело с нейронными микроконтурами, дендритными деревьями и в завершение — с нейронами. Под нейронным микроконтуром (neural micro circuit) понимается набор синапсов, организованный в шаблоны взаимосвязей, выполняющих определенную операцию. Нейронный микроконтур можно сравнить с электронным чипом, состоящим из набора транзисторов. Минимальный размер микроконтуров измеряется в микронах, а скорость операций — в миллисекундах. Нейронные микроконтурные группируются в дендритные

субблоки (dendritic subunit), составляющие дендритные деревья (dendritic tree) отдельных нейронов. Весь нейрон (neuron) имеет размеры порядка 100 микрон и содержит несколько дендритных субблоков. На следующем уровне сложности находятся локальные цепочки (local circuit) (размером порядка 1 мм), состоящие из нейронов с одинаковыми или сходными характеристиками. Эти наборы нейронов выполняют функции, характерные для отдельных областей мозга. За ними в иерархии следуют межрегиональные цепочки (interregional circuit), состоящие из траекторий, столбцов и топографических карт и объединяющие несколько областей, находящихся в разных частях мозга.

Топографические карты (topographic map) предназначены для ответа на поступающую от сенсоров информацию. Эти карты часто организуются в виде таблиц, причем визуальные, звуковые и соматосенсорные карты хранятся в стеке, сохраняющем пространственную конфигурацию конкретных точек возбуждения.

На заключительном уровне сложности топографические карты и прочие межрегиональные цепочки связываются друг с другом, образуя центральную нервную систему (central nervous system).

1.8 Модели нейронов

Единицей обработки информации в искусственной нейронной сети является нейрон. Модель такого нейрона представлена на рисунке 8. Выделяют три основных составляющих такой модели:

1. Набор синаптических связей в сети, каждая из которых характеризуется своим весом и силой. Другими словами, сигнал x_j на входе синапса j , связанного с нейроном k , умножается на вес w_{kj} . Большое значение имеет порядок индексов при обозначении веса w_{kj} . Первый индекс обозначает исследуемый нейрон, а второй – входное окончание синапса, с которым связан рассматриваемый нейрон. Кардинальным отличием искусственного нейрона от нейрона человеческого мозга является то, что знак

синаптического веса искусственного нейрона может быть как положительным, так и отрицательным.

2. Следующим элементом в искусственном нейроне является сумматор (adder). Его функция заключается в том, чтобы складывать сигналы, взвешенные относительно соответствующих синапсов нейрона. Такую операцию часто называют линейной комбинацией.

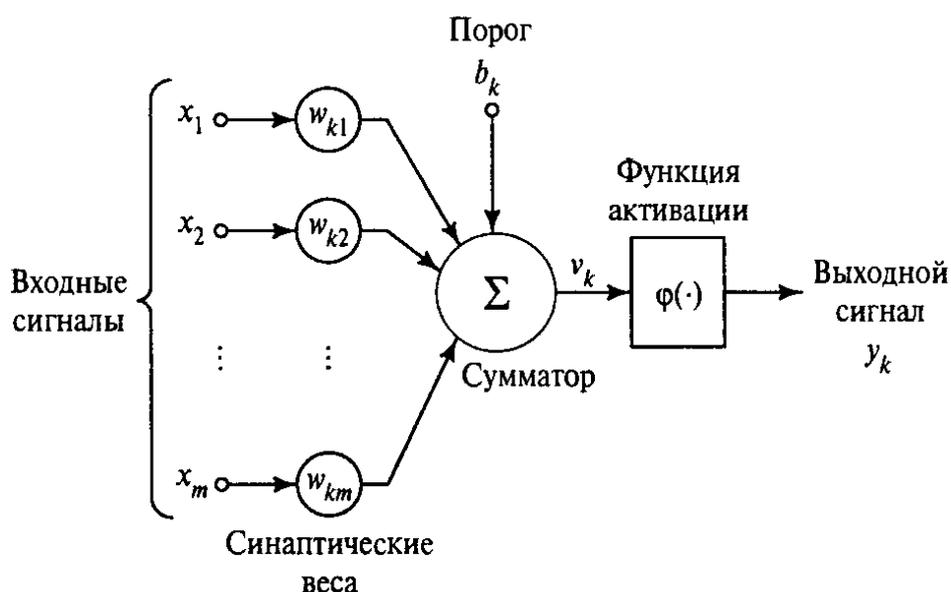


Рисунок 8 – Нелинейная модель нейрона

3. Следующим элементом нейрона является функция активации (activation function). Ее задача состоит в том, чтобы ограничивать амплитуду выходного сигнала. Другим названием этой функции является функция сжатия (squashing function). Для искусственных нейронных сетей обычным диапазоном, в котором лежит выходной сигнал, является диапазон $[0,1]$ или $[-1,1]$.

В ходе дальнейшей работы, это стало причиной того, что для прогнозирования почасового электропотребления все значения электропотребления необходимо было привести в соответствие с указанными интервалами. В противном случае, работа сети была бы некорректной и ошибка прогнозирования составляла бы значительную величину.

На рисунке 8 также можно увидеть пороговый элемент, который обозначен символом b_k . Данная величина показывает, на сколько увеличился или уменьшился входной сигнал, поступающий на функцию активации.

Функционирования нейрона k в математическом представлении можно описать следующими уравнениями:

$$u_k = \sum_{j=1}^m w_{kj} x_j,$$

$$y_k = \varphi(u_k + b_k)$$

Где x_1, x_2, \dots, x_m — входные сигналы; $w_{k1}, w_{k2}, \dots, w_{km}$ — синаптические веса нейрона k ; u_k — линейная комбинация входных воздействий (linear combiner output); b_k — порог; $\varphi(\cdot)$ — функция активации (activation function); y_k — выходной сигнал нейрона.

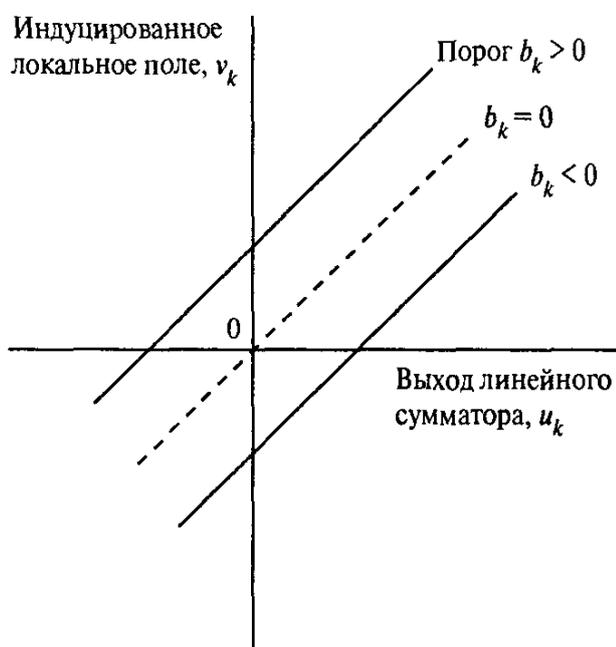


Рисунок 9 – Аффинное преобразование, вызванное наличием порога

Эффект аффинного преобразования выхода линейного сумматора u_k обеспечивается путем использования порога b_k . В модели, показанной на рисунке, постсинаптический потенциал вычисляется следующим образом:

$$v_k = u_k + b_k.$$

В частности, в зависимости от того, какое значение принимает порог b_k , положительное или отрицательное, индуцированное локальное поле (induced local field) или потенциал активации (activation potencial) v_k нейрона k изменяется так, как показано на рисунке 9.

1.9 Принципы обучения нейронных сетей

Одним из важнейших свойств искусственных нейронных сетей является их способность обучаться и в результате этого повышать свою производительность. Повышение производительности происходит со временем в результате выполнения сетью установленных правил обучения. Суть процесса обучения искусственной нейронной сети заключается в изменении синаптических весов и порогов в соответствии с парадигмой обучения. Для идеальных условий характерно то, что нейронная сеть получает информацию об окружающей среде при каждом процессе итерации.

Понятие обучения достаточно сложно в определении, так как в разных сферах деятельности это понятие включает в себя различные составляющие и соответственно под одним и тем же словом могут пониматься совершенно различные процессы. Так например процесс обучения с точки зрения школьного учителя будет рассматриваться совершенно иначе, нежели с позиции проектирования искусственной нейронной сети. С позиций нейронной сети, можно использовать следующее определение:

Обучение – это процесс настройки весовых коэффициентов и порогов таким образом, при котором обеспечивается наилучшее встраиваемость сети

в окружающую ее среду. Тип обучения определяется способом подстройки этих параметров.

Исходя из такого определения, можно записать следующую последовательность действий:

1. Поступление внешних данных на вход нейронной сети;
2. Как следствие, изменение свободных параметров искусственной нейронной сети;
3. Изменение реакции (ответа) нейронной сети на данные, поступающие на вход сети.

Данная последовательность действий называется алгоритмом обучения искусственной нейронной сети. Необходимо отметить, что невозможно создать универсальный алгоритм обучения, который бы подходил для любой архитектуры нейронной сети. Существуют лишь стратегии обучения, каждая из которых имеет свои преимущества и недостатки. Отличительной особенностью каждого алгоритма является свой уникальный способ настройки синаптических весов нейронов. Также одним из существенных параметров обучения нейронной сети, является способ ее связи с внешним миром. Иными словами, можно сказать, что алгоритмы обучения нейросетей различаются способ представления модели внешних данных, поступающих на вход сети.

На данный момент выделяют 5 основных моделей обучения:

- на основе коррекции ошибок;
- с использованием памяти;
- Хеббовское обучение;
- конкурентное обучение;
- метод Больцмана.

Обучение, основанное на коррекции ошибок, предполагает фильтрацию результатов обучения с некоторой установленной точностью. Обучение на основе памяти другими словами можно назвать, как обучение на основе обучающих данных. Метод Хебба и конкурентный подход к

обучению основаны на принципах нейробиологии. В основу метода Больцмана положены идеи статистической механики.

1.10 Обучение, основанное на коррекции ошибок

Рассмотрим простейший случай нейрона k , который является единственным вычислительным узлом выходного слоя нейронной сети прямого распространения (рисунок 10). Нейрон k функционирует под управлением вектора сигнала (signal vector) $x(n)$, который производится скрытыми слоями нейронов, которые, в свою очередь, получают данные из входного вектора (возбуждения), который передается через входной слой нейронной сети. Под n подразумевается либо дискретное время, либо, что более конкретно и чаще всего используется, — номер шага итерации определения синаптических весов нейрона k . Выходной сигнал (output signal) нейрона k обозначается $y_k(n)$. Этот сигнал является единственным выходом нейронной сети. Стоит отметить, что при рекуррентной архитектуре сети, выходные данные могут вновь подаваться в скрытый слой нейронной сети. В случае с одним выходом, сигнал будет сравниваться с желаемым выходом (desired response), обозначенным $d_k(n)$. В результате получаем сигнал ошибки (error signal) $e_k(n)$, определяемый формулой:

$$e_k(n) = d_k(n) - y_k(n).$$

Сигнал ошибки запускает последовательные процессы корректировки синаптических весов нейрона k . Эти корректировки нацелены на постепенное приближение выходного сигнала $y_k(n)$ к желаемому $d_k(n)$. Эта цель достигается за счет минимизации функции стоимости (cost function) или индекса производительности (performance index) $E(n)$, определяемой в терминах сигнала ошибки следующим образом:

$$E(n) = \frac{1}{2} e_k^2(n),$$

где $E(n)$ — текущее значение ошибки. Пошаговая корректировка синаптических весов нейрона k продолжается до тех пор, пока нейронная сеть не достигнет устойчивого состояния (steady state) (т.е. такого, при котором синаптические веса практически не изменяются). На этом процесс обучения заканчивается.

Процесс, который был описан выше называется обучением, основанным на коррекции ошибок. Снижение значений функции стоимости $E(n)$ производится по так называемому дельта-правилу, или правилу Видроу-Хоффа. Обозначим $w_{kj}(n)$ текущее значение синаптического веса нейрона k , соответствующего элементу $x_j(n)$ вектора $x(n)$, на шаге итерации n . В соответствии с дельта-правилом изменение $\Delta w_{kj}(n)$, применяемое к синаптическому весу w_{kj} на этом шаге дискретизации, задается выражением:

$$\Delta w_{kj}(n) = \eta e_k(n) x_j(n),$$

где η — некоторая положительная константа, определяющая скорость обучения (rate of learning) и используемая при переходе от одного шага процесса к другому. Эту константу обычно называют параметром скорости обучения (learning rate parameter). Дельта-правилу можно дать следующее определение - корректировка, применяемая к синаптическому весу нейрона, пропорциональна произведению сигнала ошибки на входной сигнал, его вызвавший.

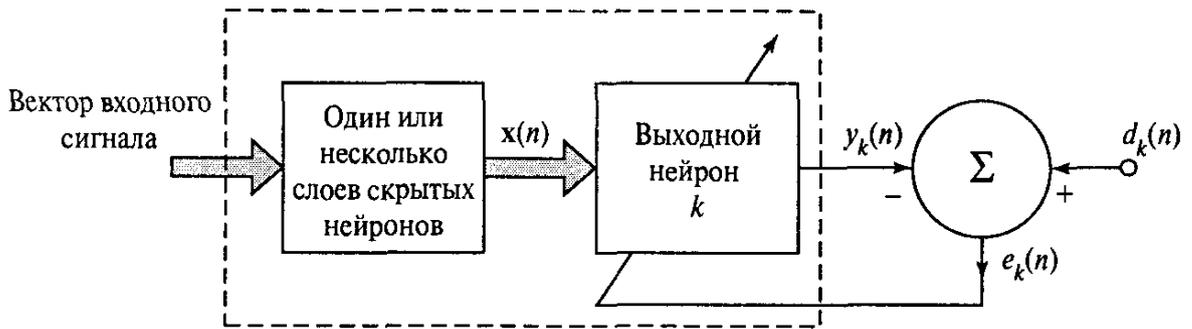


Рисунок 10 – Обучение, основанное на коррекции ошибок; блочная диаграмма нейронной сети

Из данного определения следует, что существует возможность прямого измерения величины сигнала ошибки. Такое измерение возможно только при поступлении сигнала от внешнего источника, непосредственно доступного для нейрона k . Другими словами, нейрон k должен быть «видимым» для внешнего мира (рисунок 10). На рисунке 10 можно увидеть, что обучение, основанное на коррекции ошибки, является локальным. Это означает, что корректировка синаптических весов по дельта-правилу может быть произведена в отдельном нейроне k .

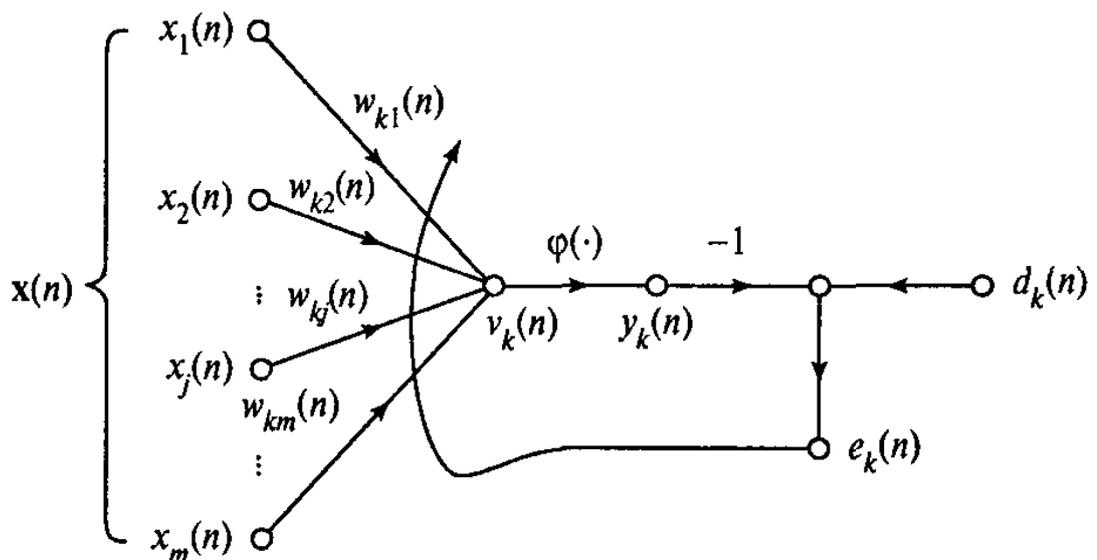


Рисунок 11 – Обучение, основанное на коррекции ошибок; граф передачи сигнала выходного нейрона

Вычислив величину изменения синаптического веса $\Delta w_{kj}(n)$, можно определить его новое значение для следующего шага итерации:

$$w_{kj}(n + 1) = w_{kj}(n) + \Delta w_{kj}(n).$$

Таким образом, $w_{kj}(n)$ и $w_{kj}(n + 1)$ можно рассматривать как изначальное и новое значение, преданное синаптическому весу w_k . В математических терминах можно записать:

$$w_{kj}(n) = z^{-1}[w_{kj}(n + 1)],$$

где z^{-1} — оператор единичной задержки (unit delay operator). Другими словами, оператор z^{-1} представляет собой элемент памяти (storage element).

На рисунке 11 представлен граф прохождения входного сигнала в процессе обучения, основанного на коррекции ошибок, для единичного нейрона k . Входной сигнал x_k и индуцированное локальное поле v_k нейрона k представлены в виде пресинаптического (presynaptic) и постсинаптического (postsynaptic) сигналов j -го синапса нейрона k . На рисунке 11 видно, что обучение на основе коррекции ошибок — это пример замкнутой системы с обратной связью (closed-loop feedback). Параметры обратной связи определяют устойчивость данной системы. В рассматриваемом случае имеется только одна обратная связь, и единственным значительным ее параметром является коэффициент скорости обучения η . Для обеспечения устойчивости или сходимости итеративного процесса обучения требуется высокая точность при выборе этого параметра. Выбор данного параметра обучения влияет на скорость обучения, а также на его точность. Другими словами, параметр скорости обучения η — это основной параметр, определяющий эффективность и производительность процесса обучения.

1.11 Обучение на основе памяти

Обучение на основе памяти предполагает хранение правильно классифицированных примеров вход-выход, что в свою очередь можно назвать памятью нейронной сети: $\{(x_i, d_i)\}_{i=1}^N$, где x_i — входной вектор, а d_i — соответствующий ему желаемый выходной сигнал. Не ограничивая общности, можно предположить, что выходной сигнал является скаляром. Например, рассмотрим задачу бинарного распознавания образов или классификации на два класса (гипотезы), C_1 и C_2 . В этом примере желаемый отклик системы d_i принимает значение 0 (или -1) для класса C_1 и значение +1 для класса C_2 . Если требуется распознать, к какому классу относится некоторый неизвестный вектор x_{test} , из базы данных выбирается выход, соответствующий входному сигналу, примерно равному x_{test} .

Все алгоритмы обучения на основе памяти включают в себя две основные составляющие:

- параметр, используемый для определения окрестности вектора x_{test} ;
- правило обучения, используемое при распознавании тестового вектора.

Каждый алгоритм отличается друг от друга именно этими двумя параметрами.

В стандартном алгоритме обучения нейронной сети на основе памяти, получившем название правила ближайшего соседа (nearest neighbor rule), в окрестность включается пример, ближайший к тестовому. Например, вектор:

$$x'_N \in \{x_1, x_2, \dots, x_N\}$$

считается ближайшим соседом вектора x_{test} если выполняется условие

$$\min_i d(x_i, x_{test}) = d(x'_N, x_{test}),$$

где $d(x'_N, x_{test})$ — Евклидово расстояние между векторами x_i и x_{test} . Класс, к которому относится ближайший сосед, считается также классом тестируемого вектора x_{test} . Это правило не зависит от распределения, используемого при генерировании примеров обучения.

При формальном исследовании правила ближайшего соседа, применяемого для решения задачи классификации образов, анализ основывается на двух следующих предположениях:

- классифицируемые примеры (x_i, d_i) независимы и равномерно распределены (independently and identically distributed) в соответствии с совместным распределением примера (x, d) .

- размерность обучающего множества N бесконечно велика.

Показано, что при этих двух предположениях вероятность ошибки классификации при использовании правила ближайшего соседа вдвое превышает байесовскую вероятность ошибки (Bayes probability error). (Байесовская вероятность ошибки — это минимальная вероятность ошибки на множестве всех правил принятия решения.) В этом контексте можно считать, что половина классификационной информации для обучающего множества бесконечного размера содержится в данных о ближайшем соседе.

Вариацией классификатора на основе ближайшего соседа является классификатор k -ближайших соседей (k-nearest neighbor classifier). Он описывается следующим образом.

Находим k классифицированных соседей, ближайших к вектору x_{test} , где k — некоторое целое число.

Вектор x_{test} относим к тому классу (гипотезе), который чаще других встречается среди k -ближайших соседей тестируемого вектора.

Таким образом, классификатор на основе k -ближайших соседей работает подобно устройству усреднения. Например, он может не учесть

единичный «выброс». Выброс (outlier) — это наблюдение, которое отличается от номинальной модели.

Стоит сказать, что описанные выше модели обучения являются наиболее приемлемыми для решения задач прогнозирования, поэтому остановимся на рассмотрении этих моделей в качестве основных для обучения ИНС для нужд энергосбытовой компании.

1.12 Парадигмы обучения нейронных сетей.

1) Обучение с учителем

Далее перейдем к рассмотрению различных парадигм обучения нейронных сетей. Первым рассмотрим парадигму обучения с учителем. На рисунке 12 представлена диаграмма, описывающая данную форму обучения. В общем случае, наличие учителя можно рассматривать как наличие знаний об окружающей среде в виде пар вход-выход. Это при условии, что внешняя среда незнакома нейронной сети. На начальном этапе учителю и обучаемой нейронной сети подается входной обучающий вектор из внешней среды. Основываясь на имеющихся знаниях, учитель передает нейронной сети желаемый отклик на данный входной обучающий вектор. Такой отклик предполагает совершение нейронной сетью оптимальных действий по достижению желаемого выхода. Другими словами, параметры сети корректируются в зависимости от входного обучающего вектора и сигнала ошибки. Сигналом ошибки называют разность между желаемым откликом и действительным выходным значением сигнала нейронной сети. Имитируя поведение учителя, искусственная нейронная сеть выполняет пошаговую корректировку всех своих параметров. Такой процесс называется эмуляцией. Этот процесс должен быть оптимальным для получения наибольшей схожести действий нейронной сети с действиями учителя. Можно сказать, что процесс обучения нейронной сети – это процесс передачи знаний учителя нейронной сети. После проведения тестирования сети и получения

удовлетворительных результатов, необходимо отключить учителя и дать возможность сети работать самостоятельно.

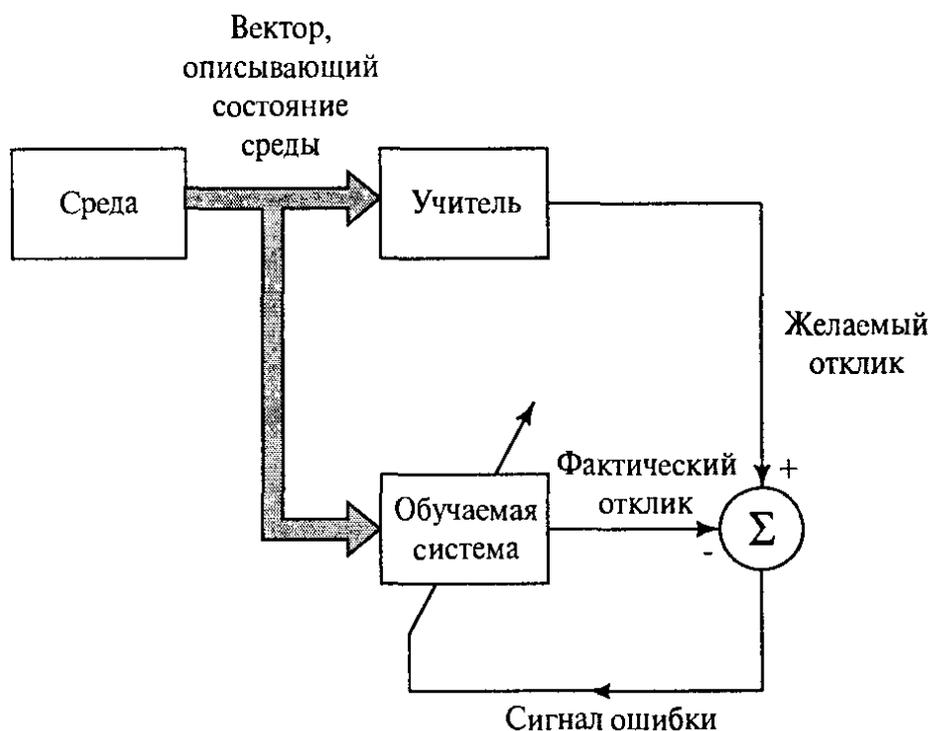


Рисунок 12 – Блочная диаграмма обучения с учителем

Такое обучение можно назвать обучением, основанным на коррекции ошибок. Данная система является системой с обратной связью, не включающей в себя окружающую среду. Эффективность работы таких систем обычно оценивают с помощью среднеквадратической ошибки или суммы квадратов ошибок. Эти значения чаще всего представляют в виде функций от свободных параметров системы. Соответственно, при наличии такой функции существует возможность построения многомерной поверхности ошибки в координатных плоскостях свободных параметров системы. Поверхность ошибки усредняется по всем возможным примерам, представленным в виде пар «вход-выход». Любое действие системы с учителем можно представить в виде конкретной точки на поверхности ошибок. Признаком повышения производительности системы является смещение точки ошибки на поверхности ошибок в сторону минимума. Этот

минимум может быть, как локальным, так и глобальным. Достижение этих минимумов возможно, если система обладает информацией о градиенте поверхности ошибок при текущем состоянии системы. Градиент поверхности ошибок — это вектор, определяющий направление, при котором изменение высоты поверхности происходит наискорейшим образом. Величина этого вектора определяет скорость изменения высоты поверхности. При обучении с учителем при помощи примеров вычисляется мгновенная оценка вектора градиента, в которой входной вектор считается функцией времени. При использовании результатов такой оценки перемещение точки по поверхности ошибок обычно имеет вид «случайного блуждания». Однако в случае наличия адекватных примеров вход-выход, достаточного времени для обучения сети, а также при использовании соответствующего алгоритма минимизации функции стоимости, система способна успешно решать задачи распознавания образов, аппроксимации функций и прогнозирования временных рядов.

2) Обучение без учителя

Описанный выше процесс обучения происходит под управлением учителя. Другой парадигмой обучения является обучение без руководителя, т.е. сеть самостоятельно осуществляет настройку весовых коэффициентов. При таком подходе отсутствуют примеры, при помощи которых бы сеть смогла достигнуть требуемых для успешного результата действий. В такой парадигме обучения выделяют два метода:

1) Обучение с подкреплением, или нейродинамическое программирование.

В обучении с подкреплением (reinforcement learning) формирование отображения входных сигналов в выходные выполняется в процессе взаимодействия с внешней средой с целью минимизации скалярного индекса производительности. На рисунке 13 показана блочная диаграмма одной из форм системы обучения с подкреплением, включающей блок «критики», который преобразовывает первичный сигнал подкрепления (primary

reinforcement signal), полученный из внешней среды, в сигнал более высокого качества, называемый эвристическим сигналом подкрепления (heuristic reinforcement signal). Оба ЭТИХ сигнала являются скалярными.

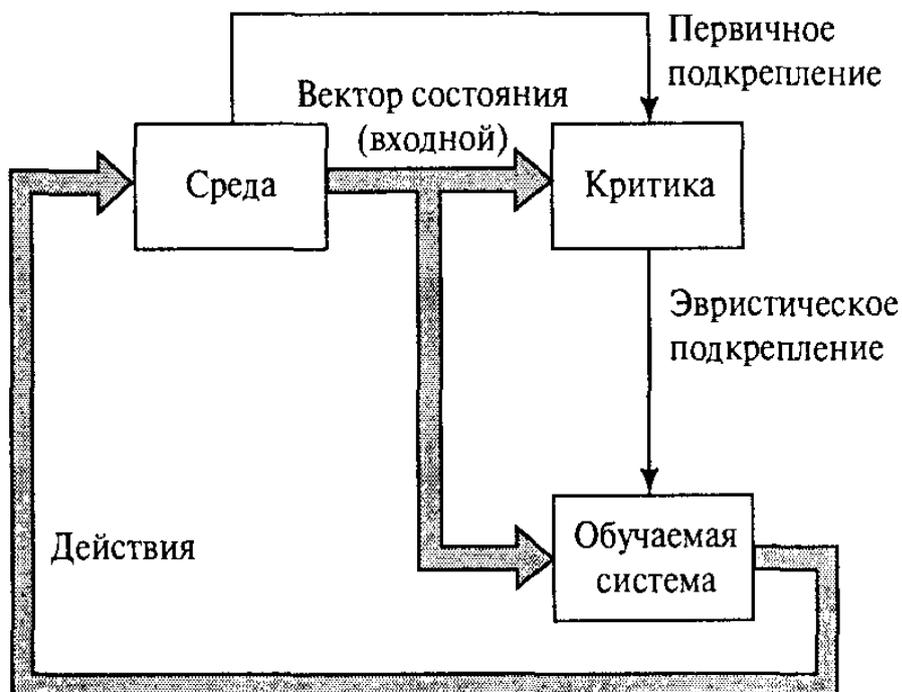


Рисунок 13 – Блочная диаграмма обучения с подкреплением

Такая система предполагает обучение с отложенным подкреплением (delayed reinforcement). Это значит, что система получает из внешней среды последовательность сигналов возбуждения (т.е. векторов состояния), которые приводят к генерации эвристического сигнала подкрепления. Целью обучения является минимизация функции стоимости перехода, определенной как математическое ожидание кумулятивной стоимости действий, предпринятых в течение нескольких шагов, а не просто текущей стоимости. Может оказаться, что некоторые предпринятые ранее в данной последовательности действия были определяющими в формировании общего поведения всей системы. Функция обучаемой машины (learning machine), составляющая второй компонент системы, определяет эти действия и

формирует на их основе сигнал обратной связи, направляемый во внешнюю среду.

Практическая реализация обучения с отложенным подкреплением осложнена по двум причинам:

- не существует учителя, формирующего желаемый отклик на каждом шаге процесса обучения;

- наличие задержки при формировании первичного сигнала подкрепления требует решения временной задачи присваивания коэффициентов доверия (temporal credit assignment). Это значит, что обучаемая машина должна быть способна присваивать коэффициенты доверия и недоверия действиям, выполненным на всех шагах, приводящих к конечному результату, в то время как первичный сигнал подкрепления формируется только на основе конечного результата.

Несмотря на эти сложности, системы обучения с отложенным подкреплением являются очень привлекательными. Они составляют базис систем, взаимодействующих с внешней средой, развивая таким образом способность самостоятельного решения возникающих задач на основе лишь собственных результатов взаимодействия со средой.

Обучение с подкреплением тесно связано с динамическим программированием (dynamic programming) — методологией, созданной Веллманом в 1957 году в контексте теории оптимального. Динамическое программирование реализует математический формализм последовательного принятия решений. Перемещая обучение с подкреплением в предметную область динамического программирования, можно взять на вооружение все результаты последнего.

Обучение без учителя не предполагает вмешательства внешнего корректора в процесс обучения и также может называться самоорганизацией искусственной нейронной сети (рисунок 14). Для сети существует только допустимый уровень ошибки, которого она должна достичь при помощи оптимизации своих свободных параметров. Обученная сеть способна

выявлять статистические закономерности входных данных и формировать внутреннее представление таким образом, чтобы при поступлении новых входных данных происходило автоматическое формирование новых классов свободных параметров сети.



Рисунок 14 – Блочная диаграмма обучения без учителя

Для подобного вида обучения пользуются правилом конкурентного обучения. Искусственная нейронная сеть состоит из двух слоев – входного и выходного. На вход сети поступают данные из внешней среды. На выходе сети нейроны конкурируют между собой за возможность вывода данных. При такой стратегии, данные способен вывести только тот нейрон, который получает наибольший отклик от входного сигнала. Таким образом в выводе данных участвует только один нейрон, а все остальные отключены.

1.13 Задачи обучения

Выбор конкретного алгоритма обучения зависит от задач, решению которых следует обучить нейронную сеть. В этом контексте можно выделить шесть основных задач, для решения которых в том или ином виде применяются нейронные сети. Перечислим эти задачи:

- ассоциативная память;
- распознавание образов;
- аппроксимация функций;
- управление;

- фильтрация;
- формирование диаграммы направленности.

Задачу создания модели краткосрочного прогнозирования электропотребления энергосбытовой компании можно сравнить с задачей управления, так как принцип решения обеих одинаков, а именно, необходимо выстроить такую модель искусственной нейронной сети и с такими весовыми коэффициентами, чтобы выходной сигнал ИНС минимально отличался от эталонного сигнала. Это приводит к высокой точности как в случае с управлением системами, так и в случае прогнозирования временных рядов.

Управление предприятием (plant) является еще одной задачей обучения, которую можно решать с использованием нейронной сети. Здесь под термином «предприятие» понимается процесс или критическая часть системы, подлежащие управлению. Ассоциация между управлением и обучением является закономерной, так как человеческий мозг является компьютером (обработчиком информации), выходами которого, как системы в целом, являются действия. В контексте задач управления сам мозг является живым доказательством возможности создания обобщенной системы управления, использующей преимущества параллельных распределенных вычислений и одновременно управляющей тысячами функциональных механизмов. Такая система является нелинейной, может обрабатывать шумы и оптимизировать свою работу в ракурсе долгосрочного планирования.

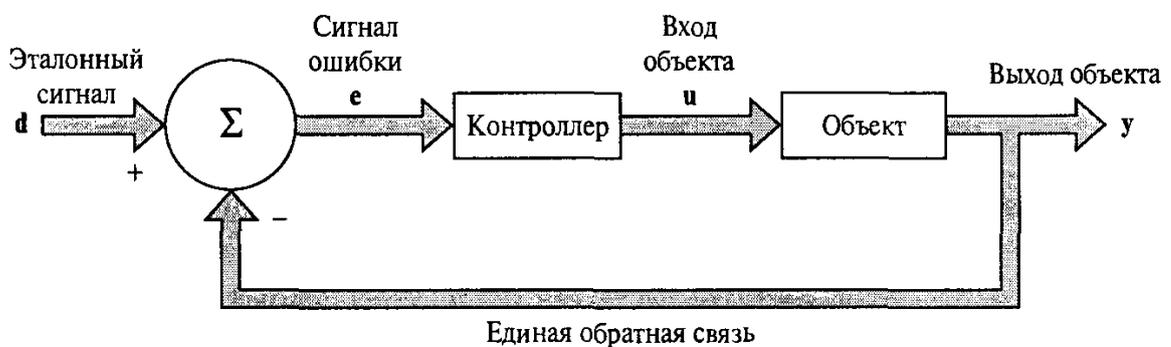


Рисунок 15 – Блочная диаграмма системы управления с обратной связью

Рассмотрим систему управления с обратной связью (feedback control system), показанную на рисунке. В данной системе присутствует только одна обратная связь, которая захватывает весь объект (т.е. выход объекта связан с его входом). Таким образом, выход объекта управления y вычитается из желаемого сигнала d , который подается из внешнего источника. В результате этого сигнал ошибки e подается на нейроконтроллер (neurocontroller) для последующей настройки свободных параметров системы. Основной задачей нейроконтроллера является поддержание такого входного сигнала, для которого выходной сигнал y максимально был бы приближен к желаемому значению d . Иначе говоря, задача нейроконтроллера заключается в инвертировании отображения сигналов вход-выход объекта управления.

На рисунке 15 сигнал ошибки e распространяется через нейроконтроллер, прежде чем достигнет объекта управления. Следовательно, для настройки свободных параметров объекта управления в соответствии с алгоритмом обучения на основе коррекции ошибок необходимо знать матрицу Якобиана:

$$J = \left\{ \frac{\partial y_k}{\partial u_j} \right\},$$

где y_k — элемент выходного сигнала y объекта управления; u_j — элемент вектора входов объекта и. К сожалению, частные производные $\partial y_k / \partial u_j$ для различных k и j зависят от рабочей точки объекта управления и, следовательно, не известны. Для их оценки можно использовать два следующих подхода:

- не прямое обучение (indirect learning). С использованием текущих измерений входа и выхода объекта управления строится модель нейронной сети, воспроизводящая эту зависимость. Эта модель, в свою очередь, используется для оценки матрицы Якобиана J . Частные производные, составляющие эту матрицу, впоследствии используются в алгоритме

обучения на основе коррекции ошибки для настройки свободных параметров нейроконтроллера;

-прямое обучение (direct learning). Знаки частных производных $\partial u_k / \partial u_j$ в общем случае известны и обычно остаются постоянными в некотором динамическом диапазоне значений объекта управления. Это наводит на мысль, что частные производные можно аппроксимировать по их знакам. Абсолютные значения частных производных задаются распределенным представлением в свободных параметрах нейроконтроллера. Таким образом, свободные параметры нейроконтроллера можно настроить непосредственно с помощью объекта управления.

1.14 Обучающие данные для нейронной сети

Обучение нейронной сети в большинстве случаев предполагает наличие обучающих данных. По возможности данные должны отвечать требованиям представительности, случайности или последовательности. Все зависит от типа решаемой задачи. В нашем случае для прогнозирования обучающие данные представляют собой временной ряд с указанием величин, которые желательно было бы получить на выходе нейронной сети. В этом случае обучение можно назвать контролируемым: исходные данные подаются учителем на вход нейронной сети, а полученные выходные данные сравниваются с эталонными.

Такой вид обучения можно рассматривать как решение задачи оптимизации, целью которого является минимизация функции ошибок при данном множестве примеров. Это достигается путем настройки весовых коэффициентов сети. Достижение минимума функции ошибки называется сходимостью процесса обучения. Возможность такой сходимости доказал Розенблат.

Поскольку зависимость ошибки от весовых коэффициентов сети является нелинейной, то невозможно получить функцию данной зависимости и для поиска минимальной функции ошибки приходится прибегать к

итерационному процессу. Другими словами, необходимо повторять обучающий алгоритм на конкретных примерах входных данных. На данный момент существует огромное множество обучающих алгоритмов, которые отличаются между собой способами оптимизации и критериями оценки ошибки обучения. Обычно в качестве меры погрешности обучения принимается средняя квадратичная ошибка (СКО):

$$E = \sqrt{\frac{\sum_{i=1}^M (d_i - y_i)^2}{M}},$$

где M – число примеров в обучающем множестве.

Уменьшение величины E производится с помощью поиска градиента поверхности ошибки. Изменение весов происходит в направлении обратном к направлению наибольшей крутизны градиента для функции:

$$W(t + 1) = W(t) + \varepsilon \cdot \frac{\partial E}{\partial W}$$

Здесь ε - определяемый исследователем параметр, который называется коэффициентом обучения.

Способы обеспечения процесса сходимости и его ускорения:

1. Выбор начальных весовых коэффициентов

Для ускорения процесса сходимости и повышения его эффективности возможно применение заранее рассчитанных весовых коэффициентов, исходя из характера данных и задачи обучения. Чаще всего при проектировании искусственных нейронных сетей используется подход к решению данной проблемы, при котором всем весовым коэффициентам придается минимальное и примерно равное между собой значение. Но такой способ подготовки весовых коэффициентов не может обеспечить уменьшение времени сходимости, а также достижения глобального минимума функции ошибки.

2. Упорядочение данных

При решении задач аппроксимации функций и распознавания образов необходимо перемешивать входные данные, поступающие в нейронную сеть, случайным образом. Это обязательно делать в связи с тем, что при решении задачи обучения, сеть может «запомнить» последовательность данных и принимать это как истинное решение задачи. Это недопустимо, поскольку в таком случае сеть будет обучаться неверно и в последствии не сможет правильно работать с тестовой выборкой данных. Стоит также отметить, что перемешивание данных в случайном порядке недопустимо в случае решения задачи прогнозирования временных рядов.

3. Импульс смещения

При изменении весовых коэффициентов иногда прибегают к суммированию изменений весов на текущем шаге итерации и на предыдущем шаге. Это позволяет учитывать данные, полученные при предыдущем обучении, что в теории должно повышать скорость и эффективность процесса сходимости. В этом случае формула изменения веса синаптической связи будет выглядеть следующим образом:

$$\Delta W_{ij}(t+1) = \mu * \Delta W_{ij}(t) - (1 - \mu)\varepsilon \frac{\partial E}{\partial W_{ij}},$$

где μ - число в интервале (0,1), которое задается исследователем.

4. Управление величиной шага сети

ε - величина шага сети. Другими словами, это - критерий точности обучения сети. Чем больше его величина, тем большей будет корректировка весового коэффициента на следующем шаге итерации. В случае его малой величины, корректировка соответственно будет меньше и как следствие увеличится время обучения сети. Однако, в таком случае возрастают шансы на попадание в окрестность минимума ошибки. Поэтому данный параметр является очень важным процессе обучения нейронной сети. В современных

программных комплексах существует возможность самостоятельного выбора пользователем величины шага сети. Также зачастую применяется линейная или экспоненциальная зависимость величины шага сети от количества итераций.

5. Оптимизация архитектуры сети

Сложной проблемой при проектировании нейронных сетей является выбор изначального количества скрытых слоев сети и числа нейронов в них. Если нейронов будет недостаточно, то сеть будет постоянно терять данные, в следствии чего не сможет обработать объем информации, необходимый для решения поставленной задачи. Если нейронов будет слишком много, то сеть «выучит» имеющиеся обучающие примеры и это приведет к тому, что она не научится распознавать образы и закономерности в данных и будет действовать строго по заученному шаблону.

Для решения этой проблемы применяется 2 основных подхода:

- деструктивный подход - проектируется сеть заведомо большего размера, чем нужно, и в процессе обучения из нее удаляются лишние связи и нейроны;

- конструктивный подход – к маленькой сети в зависимости от результатов тестирования добавляются необходимые элементы.

Из теоремы об отображении практически любой функции с помощью многослойной нейросети следует, что обучаемая нейронная сеть способна сама подстроиться под любые данные с целью минимизации суммарной квадратичной ошибки. Чтобы этого не происходило, при обучении нейросетей используют следующий способ проверки сети. Для этого делается разбивка входных данных на обучающую и тестовую выборки. Обучающую используют для обучения сети и изменения ее весовых коэффициентов. Тестовую в свою очередь используют только для проверки результатов обучения. Весовые коэффициенты сети при этом не меняются. Обучение происходит в правильном направлении, если уменьшение среднеквадратического отклонения выходных данных от желаемых

происходит как на обучающей, так и на тестовой выборке. Иначе может снижаться ошибка на обучающей выборке, но происходит ее увеличение на тестовой. Последнее означает, что сеть "переобучилась" и уже не может быть использована для прогнозирования или классификации. В этом случае немного изменяются веса нейронов, чтобы вывести сеть из окрестности локального минимума ошибки.

Проведя анализ двух различных методов прогнозирования планового почасового электропотребления, можно сделать вывод о том, что применение для прогноза искусственных нейронных сетей позволяет устранить недостатки статистического метода и добиться высоких показателей в точности прогноза, повысить автоматизацию процесса прогнозирования, сократить время составления прогноза и исключить субъективную оценку эксперта.

К проблемам использования нейронных сетей при прогнозировании электропотребления можно отнести большую трудоемкость процесса подготовки базы данных погодных и иных внешних условий, влияющих на конечную величину потребления электроэнергии, а также сложность процесса обучения нейронной сети и последующей ее проверки на тестовой выборке данных.

Однако, несмотря на все возможные трудности при создании нейросети, преимущества ее перед статистическим методом прогнозирования очевидны, что делает ее создание целесообразным и выгодным для компании, участвующей в торговле на ОРЭМ. Также возможно совмещение работы эксперта и искусственной нейросети, так как не редко складываются нестандартные ситуации, в которых для принятия верного решения недостаточно статистических данных, а необходим опыт и практика, которыми обладает только эксперт-аналитик.

Вывод по главе 1

В главе были рассмотрены различные методы прогнозирования электропотребления энергосбытовой компании, проведен их анализ и сравнение. Результатом этого стал вывод о том, что среди всех методов прогнозирования, в том числе методы экспертной оценки (эвристический метод), статистические методы и методы, основанные на работе искусственных нейронных сетей, наиболее перспективными выглядит последний метод. Это продиктовано рядом его преимуществ перед остальными методами, а именно:

- высокая степень автоматизации процесса прогнозирования;
- высокая точность прогнозов;
- сравнительно небольшое количество времени, требуемое для составления прогноза;
- низкая степень зависимости от субъективного мнения эксперта.

2 Разработка модели прогнозирования на базе искусственной нейронной сети

Для реализации выбранной методики прогнозирования, необходимо решить следующие основные задачи: выбор программного продукта, на базе которого будет построена ИНС, выбор архитектуры нейронной сети из числа подходящих для решения задач прогнозирования, выбор метода обучения нейронной сети, также адекватного поставленной задаче.

Каждый из этих пунктов может иметь несколько возможных вариантов решений. Далее рассмотрим наиболее предпочтительные из возможных методов построения ИНС и выберем оптимальный для прогнозирования электропотребления энергосбытовой компании.

2.1 Выбор программного продукта для реализации модели прогнозирования на базе искусственной нейронной сети

Для решения задач прогнозирования электропотребления в электроэнергетике широко применяются различные программные комплексы, основанные на статистических и нейросетевых методах прогнозирования временных рядов. Существует большое количество программных комплексов, позволяющих составлять прогнозные модели на базе искусственных нейронных сетей (ИНС). Одним из таких программных продуктов является STATISTICA.

STATISTICA Automated Neural Networks – программный комплекс, разработанный компанией StatSoft для решения разнообразных задач, в том числе и прогнозирования временных рядов, с использованием искусственных нейронных сетей с различной архитектурой и алгоритмами обучения.

К преимуществам данного программного пакета можно отнести:

- полностью русифицированный пользовательский интерфейс;

- высокий уровень визуализации интерфейса, а также простота и понятность процесса создания, обучения и тестирования ИНС, не требующий от исследователя серьезных знаний теории ИНС;

- наличие функции *Автоматизированная нейронная сеть (АНС)* позволяет выбрать оптимальный вариант сети, исходя из конкретной поставленной задачи, избегая при этом длительного поиска наилучшей модели путем «проб и ошибок»;

- возможность создания ансамблей нейронных сетей (соединение нескольких ИНС) неограниченного размера;

- возможность генерации исходного кода на языках C, C++, C#, Java, PMML (Predictive Model Markup Language), который может быть легко интегрирован во внешнюю среду для создания собственных приложений.

Однако при всех вышеперечисленных достоинствах, данный программный комплекс имеет существенный недостаток, а именно высокую стоимость как самого программного обеспечения, так и обучения работе с данным комплексом.

В этой связи был рассмотрен следующий программный комплекс, разработанный компанией MathWorks, а именно пакет прикладных программ MATLAB. В состав MATLAB входит модуль Neural Network Toolbox для создания и обучения ИНС.

К преимуществам данного программного пакета можно отнести:

- широкий выбор различных видов архитектур ИНС, а также алгоритмов обучения, что позволяет создать наиболее адекватную нейронную сеть, исходя из поставленных практических задач;

- возможность соединения различных нейронных сетей между собой, что позволяет создавать ИНС, учитывающую влияние нескольких факторов на конечный выходной параметр;

- возможность генерации исходного кода искусственной нейронной сети;

- доступность программного комплекса для широкого круга пользователей.

Также для решения задачи построения ИНС был рассмотрен программный комплекс FANN (Fast Artificial Neural Network) — открытое программное обеспечение для построения искусственных нейронных сетей. Для работы в данной программной среде могут быть использованы до 20 языков программирования, включая C#, Java, C++, PHP.

Главным преимуществом комплекса является его доступность, однако для создания различных моделей ИНС от пользователя программы требуется владение языками программирования на достаточно высоком уровне, что значительно усложняет процесс построения, обучения и тестирования ИНС.

Для составления модели прогнозирования на базе ИНС был выбран программный продукт MATLAB. Такой выбор обусловлен наличием опыта работы в MATLAB и его доступностью, а также большими вычислительными возможностями данного программного комплекса.

2.2 Составление базы данных почасового электропотребления

Необходимым условием адекватной работы ИНС является наличие как можно более глубокой ретроспективной базы данных о почасовом электропотреблении. Это необходимо для выявления закономерностей изменения графика электропотребления как в краткосрочной, так и в долгосрочной перспективе, поскольку форма графиков электропотребления в различных месяцах значительно отличается друг от друга. Также это объясняется еще и тем, что особенностью электропотребления энергосбытовой компании является значительная доля населения во всем объеме потребления (50% и более). В результате этого возрастает зависимость электропотребления от погодных условий (температуры, освещенности, наличия или отсутствия осадков) и продолжительности светового дня. Минимальным требованием к размерам ретроспективной базы является наличие данных об электропотреблении за 3 года. Однако наличие

более обширной базы данных будет только способствовать увеличению точности прогноза. Также необходимо отметить, что описание погодных условий должно быть как можно более подробным за каждый час суток, информация о которых содержится в базе. При составлении модели ИНС была доступна база данных почасового электропотребления с 2010 по 2015 год включительно. Такой объем данных является достаточным для обучения и последующей адекватной работы сети на тестовых выборках данных.

2.3 Адаптация входных данных для обучения искусственной нейронной сети

Особенностью работы ИНС, созданных в среде MATLAB, является то, что матрица входных значений для обучения сети должна иметь не менее 100 значений в одной строке. В связи с этим была проведена адаптация обучающих данных к требованиям программы. Задача создания ИНС для краткосрочного прогнозирования могла быть решена только при условии разбиения часовых значений электропотребления на получасовые, т.к. при использовании часовых значений, прогноз мог быть составлен минимум для четырех суток, что не всегда является удобным и необходимым. Поэтому была использована база данных автоматизированной информационно измерительной системы коммерческого учета электроэнергии (АИИС КУЭ), которая содержала все необходимые данные за период с 2010 по 2015 год в формате, необходимом для решения поставленной задачи.

Пробные испытания ИНС с различной архитектурой и информацией, полученной из руководства пользователя MATLAB Neural Network Toolbox, показали, что для наиболее точной работы нейронной сети необходимо все имеющиеся данные о получасовом электропотреблении привести к виду: $P=[0;1]$, где P – значения получасового электропотребления.

2.4 Выбор алгоритма обучения нейронной сети

Для обучения ИНС были выбраны два алгоритма, которые наиболее часто используются при прогнозировании временных рядов.

Первым использовался алгоритм Левенберга-Марквардта, который основан на достижении наименьшей среднеквадратической ошибки. Обучение сети прерывается в тот момент, когда перестает происходить ее уменьшение. К достоинствам данного алгоритма обучения ИНС можно отнести быстроту обучения и достаточно низкую среднеквадратическую ошибку, которая составляет порядка $3,35E-05$. Однако, при использовании этого алгоритма было выявлено, что ошибка на тестовой выборке была выше, чем на тренировочной. Поэтому данный алгоритм обучения ИНС не был выбран для решения поставленной задачи.

Второй алгоритм обучения ИНС – алгоритм основанный на регуляризации Байеса. Суть данного алгоритма заключается в том, что изменение весов синаптических функций прекращается при достижении наименьшей среднеквадратической ошибки. Обучение по данному алгоритму занимает больше времени по сравнению с алгоритмом Левенберга-Марквардта, но при этом достигается минимальная среднеквадратическая ошибка (порядка $2,02E-05$). Также при прогнозировании временных рядов, ошибка на тестовой выборке становится меньше, чем на обучающей.

Эти факты обуславливают выбор данного метода для дальнейшего построения модели ИНС.

2.5 Выбор архитектуры нейронной сети

На начальном этапе исследования для прогнозирования электропотребления использовалась сеть прямого распространения данных и обратного распространения ошибки (Feed-forward backprop). На рисунке 16 показана схема данной сети. При такой архитектуре сети для ее обучения необходимо задать целевые значения временного ряда, которые являются эталоном при составлении прогноза и входные данные, которые подлежат

обработке. После обучения, ИНС подбирает весовые коэффициенты таким образом, чтобы среднеквадратическое отклонение выходных значений от эталона было минимальным. К преимуществам такой архитектуры можно отнести быстроту работы сети.

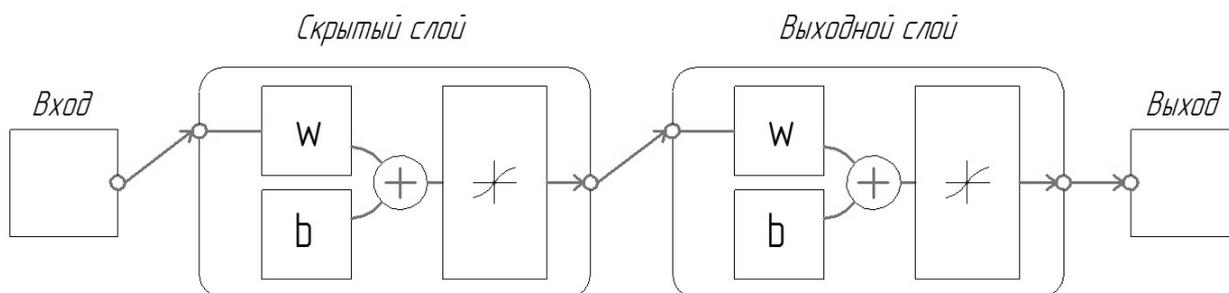


Рисунок 16 – Схема ИНС прямого распространения данных и обратного распространения ошибки (Feed-forward backprop)

Недостатками такой сети является низкая точность и невозможность использования для повторного обучения весовых коэффициентов, полученных при первичном обучении сети. Иными словами, сеть обучается каждый раз «с нуля», что не способствует увеличению точности, а напротив, приводит к различным решениям при одном и том же наборе данных.

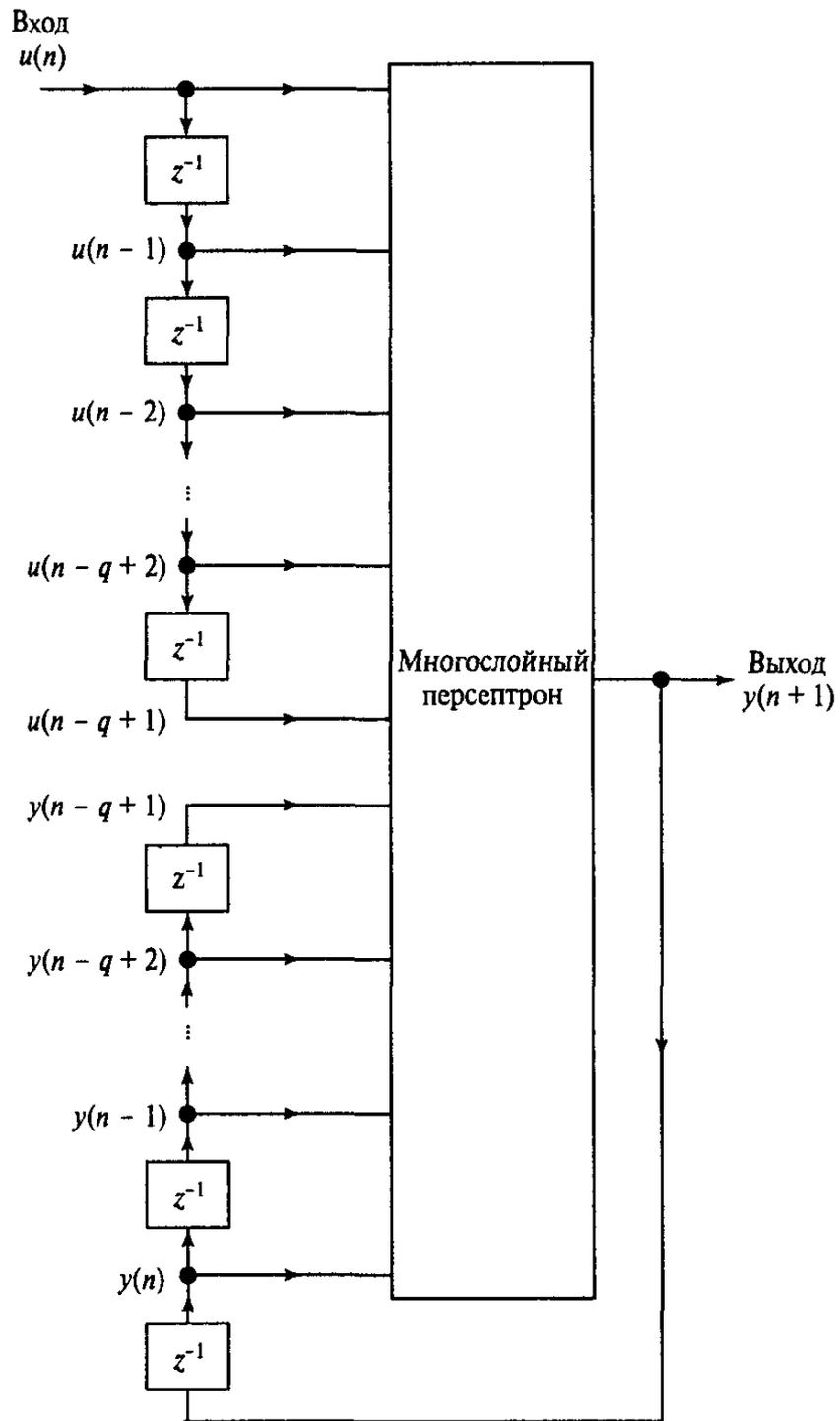


Рисунок 17 – Модель нелинейной авторегрессии с внешними входами (NARX)

На рисунке 17 показана архитектура обобщенной рекуррентной сети, построенной на базе многослойного персептрона. Эта модель имеет единственный вход, который применяется к памяти на линиях задержки, состоящей из q элементов. Она имеет единственный выход, замкнутый на вход

через память на линиях задержки, которая также состоит из q элементов. Содержимое этих двух блоков памяти используется для питания входного слоя персептрона. Вход модели обозначается как $u(n)$, а соответствующий выход — $y(n+1)$. Это значит, что выход модели упреждает ее вход на одну единицу времени. Таким образом, вектор сигнала, подаваемый на вход персептрона, состоит из окна данных, состоящего из следующих элементов:

- текущее и предыдущее значения входного сигнала: $u(n), u(n-1), \dots, u(n-q+1)$, которые представляют сети, имеющие внешнее происхождение;
- значения выходного сигнала $y(n), y(n-1), \dots, y(n-q+1)$ в предшествующие моменты времени, от которых зависит выход модели $y(n+1)$.

Таким образом, рекуррентную сеть, показанную на рисунке, можно рассматривать как модель нелинейной авторегрессии с внешними входами (nonlinear autoregressive with exogenous inputs model - NARX). Динамика модели NARX описывается следующим образом:

$$y(n+1) = F(y(n), \dots, y(n-q+1), u(n), \dots, u(n-q+1)),$$

где F — некоторая нелинейная функция своих аргументов. Обратите внимание, на рисунке предполагается, что обе памяти на дискретной линии задержки имеют размер q . В общем случае эти размеры могут отличаться.

Общий вид искусственной нейронной сети нелинейной авторегрессии представлен на рисунке 18. При такой архитектуре сети необходимо задавать как входные значения функции, так и целевые. Данные на входе, проходя через сеть, обрабатываются в соответствии с весовыми коэффициентами, а затем снова попадают на вход сети, тем самым заменяя собой обратное распространение ошибки. Это дает возможность применения полученных после первичного обучения весовых коэффициентов при повторном обучении сети, что в свою очередь повышает точность ее работы. Также при

такой архитектуре ошибка в тестовой выборке ниже, чем в обучающей, что говорит о правильном обучении сети.

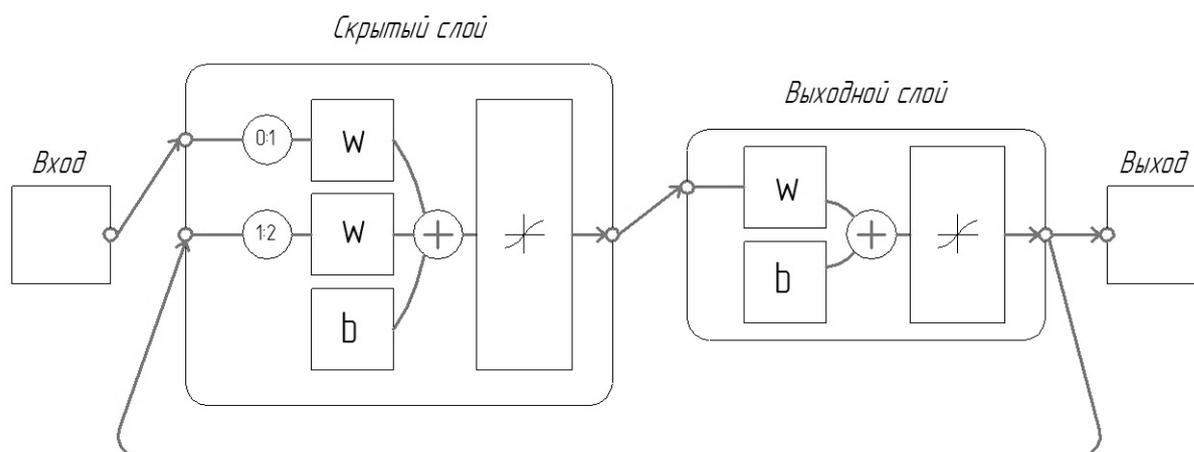


Рисунок 18 – Схема ИНС нелинейной авторегрессии с внешними входами (NARX)

2.6 Сравнение результатов прогнозирования при различных моделях построения ИНС

На первом этапе исследования для моделирования процесса электропотребления была выбрана сеть, обученная по алгоритму Левенберга-Марквардта и имеющая архитектуру прямого распространения данных и обратного распространения ошибки (Feed-forward backprop).

Прогноз составлялся для 25-26 февраля 2016 года. Данные для обучения сети были взяты из аналогичных периодов предыдущих лет. Для наглядности, при испытании новой модели ИНС использовался один и тот же набор данных, чтобы можно было сравнить показатели качества прогнозирования каждой из моделей. На рисунке 19 представлен график полученный с использованием алгоритма обучения Левенберга-Марквардта и сети с прямым распространение данных и обратным распространением ошибки.

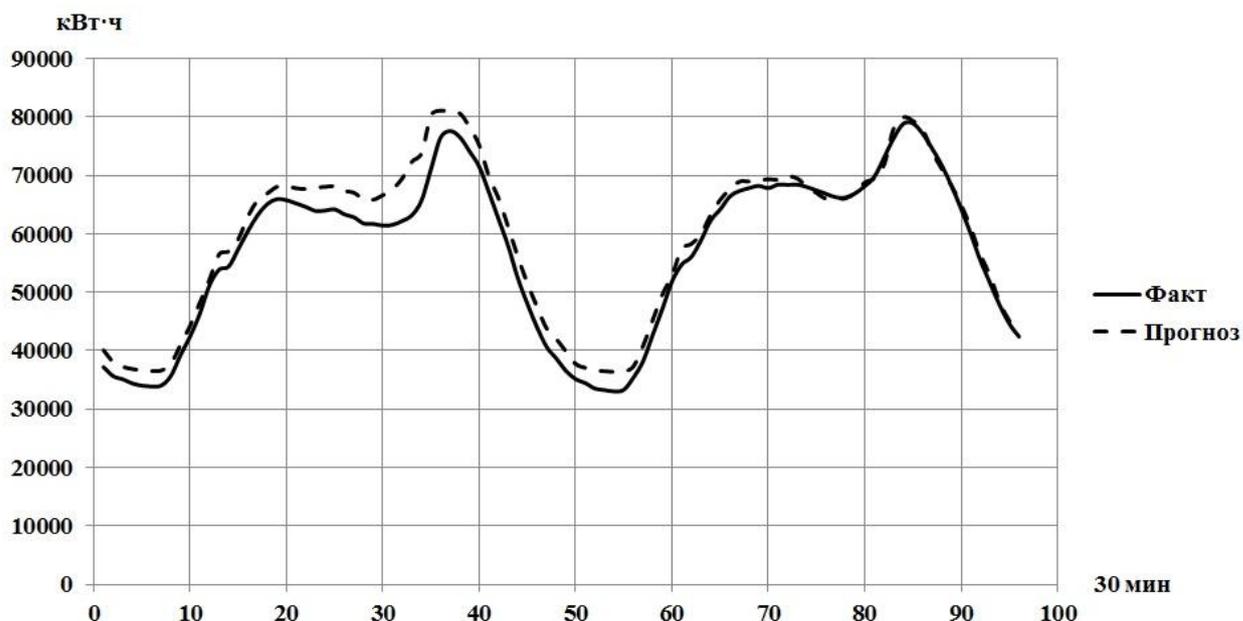


Рисунок 19 – Фактический и прогнозный графики электропотребления энергосбытовой компании для 25-26 февраля 2016г. (алгоритм Левенберга-Марквардта, архитектура Feed-forward backprop)

Как видно из графиков, прогноз на 25 февраля составлен менее точно, чем на 26 февраля. Это объясняется тем, что разброс в данных за аналогичный период в первом случае был больше, чем во втором, т.е. в базе данных электропотребления графики, использовавшиеся для составления прогноза на 25 февраля, значительно отличались друг от друга. Это объясняется различием погодных условий в сутках, использовавшихся для прогноза. Однако стоит отметить, что не смотря на это, искусственная нейронная сеть в целом правильно распознала закономерности построения графика и результатом этого стало построение похожего графика, хотя и точность его построения оказалась невысока – 6,32% отклонений по модулю от суммарного объема за сутки, что в количественном выражении составило 170 006 кВт·ч. На первый взгляд может показаться, что при объеме суточного электропотребления в зимний период, равном примерно 2 700 000 кВт·ч, такое отклонение является незначительным, однако, учитывая то, что любое отклонение прогнозного графика от фактического оплачивается на балансирующем рынке, такое отклонение не является приемлемым. В этом

случае, компания понесет слишком большие затраты на перепродажу электроэнергии на балансирующем рынке.

Рассмотрим результаты прогнозирования на 26 февраля. Выбранная модель нейронной сети показала достаточно хорошие результаты прогнозирования. Ошибка составила 2,33% от общего объема электропотребления (покупка 0,15%, продажа 2,18%). В абсолютном выражении она равна 65 161 кВт·ч отклонений по модулю за операционные сутки. Такой результат является приемлемым для успешной работы компании на ОРЭ(М).

На следующем этапе исследования была опробована модель нейронной сети, в которой использовалась та же архитектура, но иной принцип обучения – регуляризация Байеса. На рисунке 20 представлен полученный с помощью данной сети прогнозный график электропотребления.

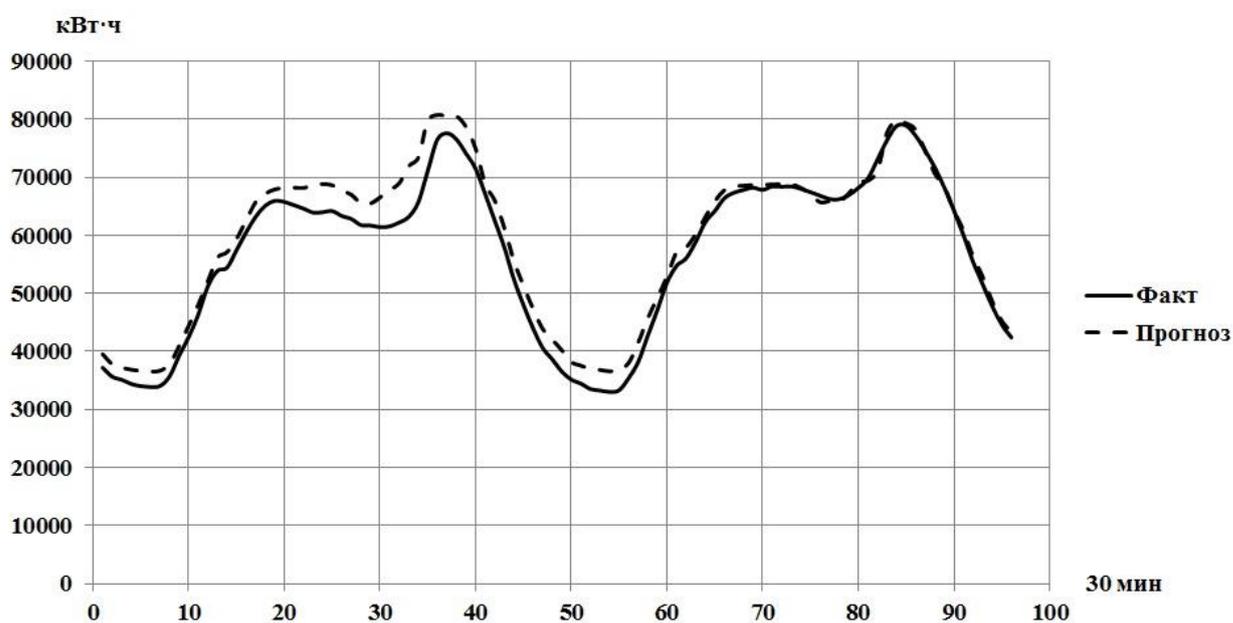


Рисунок 20 - Фактический и прогнозный графики электропотребления энергосбытовой компании для 25-26 февраля 2016г. (регуляризация Байеса, архитектура Feed-forward backprop)

Графики схожи с предыдущими, однако, стоит отметить, что точность составления прогноза в данном случае несколько ниже, чем в предыдущем –

она составила для 26 февраля 2,36% по модулю за операционные сутки. В абсолютном выражении это 66 083 кВт·ч.

Далее рассмотрим нейронную сеть с иной архитектурой и также опробуем на ней те же способы обучения. Искусственная нейронная сеть нелинейной авторегрессии с внешними входами (nonlinear autoregressive with exogenous inputs model - NARX) имеет следующую особенность: при такой архитектуре, обучение происходит в 2 этапа. На первом этапе происходит первичное определение весовых коэффициентов синаптических функций. Это делается до тех пор, пока ошибка на тестовой выборке данных не станет меньше, чем ошибка на тренировочной выборке. После того, как это условие выполнено, наступает второй этап обучения, на котором полученные ранее весовые коэффициенты уточняются путем повторного обучения сети, но уже с учетом этих коэффициентов. В результате достигается минимизация отклонения прогнозных значений от фактических. На рисунке 21 представлен график электропотребления, полученный с помощью нейронной сети с нелинейной авторегрессией и обучением Левенберга-Марквардта.

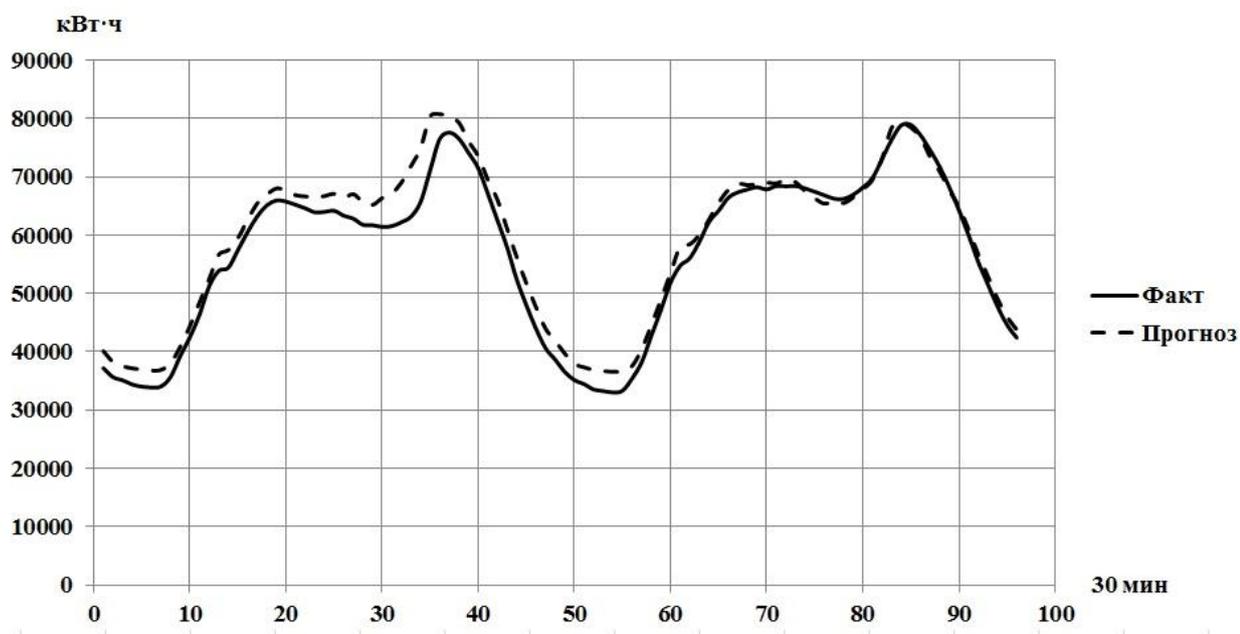


Рисунок 21 - Фактический и прогнозный графики электропотребления энергосбытовой компании для 25-26 февраля 2016г. (обучение Левенберга-Марквардта, архитектура NARX)

Результаты прогнозирования показали некоторое увеличение точности, но только в прогнозе на 25 февраля. Суммарное отклонение по модулю в процентном соотношении от общего электропотребления за сутки составило 5,88% или 158 156 кВт·ч. Такой результат нельзя признать удовлетворительным, однако стоит отметить, что особенности архитектуры сети и возможность уточнения весовых коэффициентов позволили достичь снижения ошибки. Это можно рассматривать как положительную сторону такого подхода к построению модели искусственной нейронной сети.

Что же касается 26 февраля, то здесь мы наблюдаем увеличение ошибки прогнозирования. В процентном выражении она составила 2,43%, а в абсолютном - 67 883 кВт·ч. Такие результаты объяснимы тем, что благодаря возможности уточнения весовых коэффициентов, ИНС лучше работает с набором данных, в которых существует больший разброс, как в случае с 25 февраля. В аналогичных периодах, бравшихся в качестве входных данных для прогнозирования на 26 февраля, наблюдаются более схожие графики электропотребления, в результате чего нейронная сеть «переобучилась», излишне скорректировав весовые коэффициенты синаптических функций. Это и привело к увеличению ошибки прогнозирования.

Последним вариантом построения модели нейронной сети является совмещение архитектуры NARX и обучения по принципу регуляризации Байеса. Такая модель должна была стать наиболее перспективной для дальнейшей работы, т.к. и архитектура сети и принцип обучения направлены на поиск весовых коэффициентов, позволяющих с наибольшей точностью осуществлять прогнозирование временных рядов. На рисунке 22 представлен график, составленный при помощи данной модели искусственной нейронной сети. Также как и в предыдущих вариантах прогноз составлялся для 25-26 февраля 2016г.

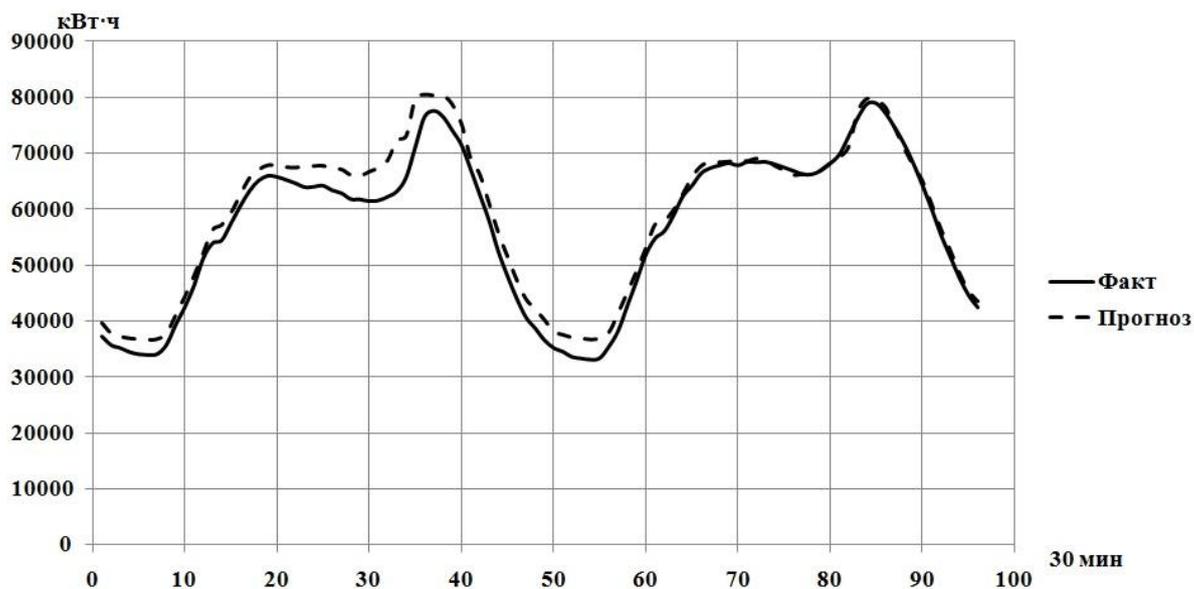


Рисунок 22 - Фактический и прогнозный графики электропотребления энергосбытовой компании для 25-26 февраля 2016г. (регуляризация Байеса, архитектура NARX)

Как видно из рисунка 22, принципиальных различий с предыдущими прогнозными графиками нет, однако, в данном случае достигнута максимальная точность прогнозирования за 26 февраля. Ошибка составила 2,27% от общей суммы электропотребления за сутки или 63 755 кВт·ч в абсолютном выражении. Этот результат является очень хорошим даже в сравнении с прогнозированием при участии в работе эксперта-аналитика. Можно говорить о том, что такой вариант модели искусственной нейронной сети является оптимальным, так как подобные результаты были получены и на аналогичных выборках данных.

Вывод по главе 2

В главе были рассмотрены вопросы реализации искусственной нейронной сети на программном уровне. Для этого был выбран программный комплекс Matlab. Также произведен выбор архитектуры сети и метода ее обучения. Проведено тестирование различных вариантов построения сети. Проведен анализ результатов прогнозирования при каждом из вариантов. Выбран наиболее эффективный способ построения сети. Тестирование показало, что наилучших результатов по точности прогноза достигает сеть с нелинейной авторегрессией и принципом обучения Байеса. Эта сеть выбрана для дальнейшего исследования.

3 Введение в модель прогнозирования сведений о метеофакторах, влияющих на формирование графика электропотребления

3.1 Оценка эффективности прогнозирования выбранной модели нейронной сети на различных типах данных

Для проверки правильности построения сети необходимо было провести испытания на различных типах графиков. Так как в разные времена года, графики электропотребления отличаются значительно, то такая проверка покажет, насколько верно искусственная нейронная сеть распознает закономерности в рядах данных и способна перестраиваться в соответствии с изменением характера данных.

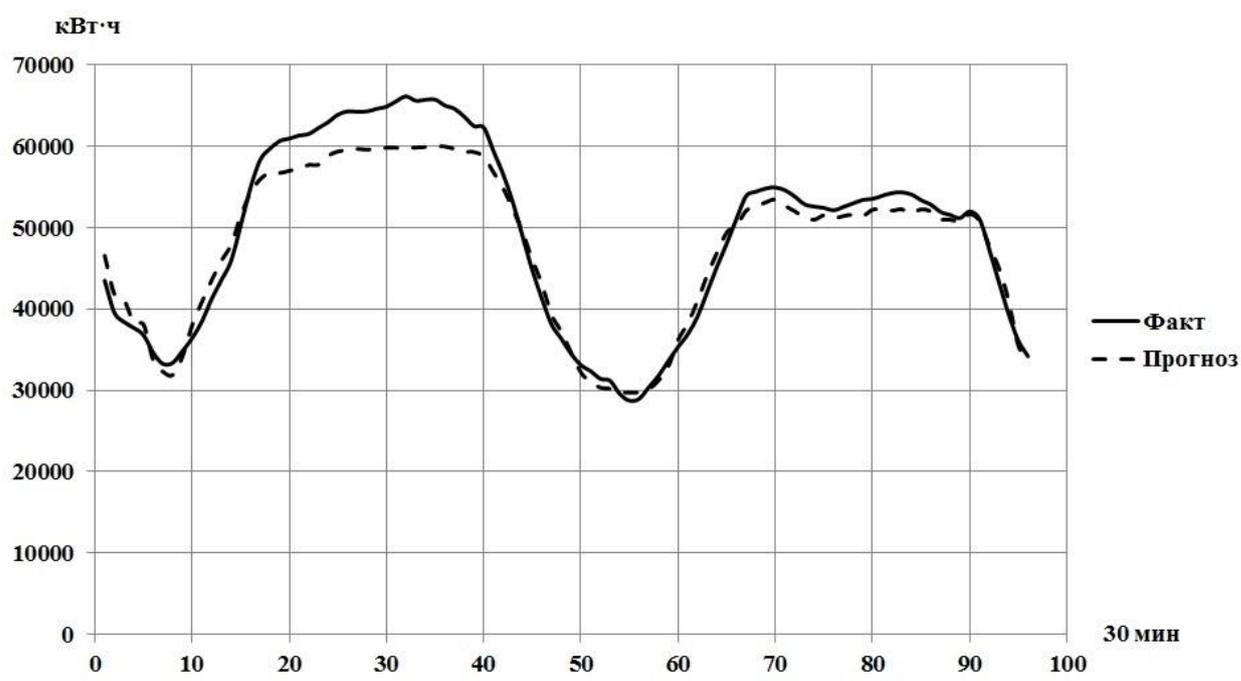


Рисунок 23 - Фактический и прогнозный графики электропотребления энергосбытовой компании для 29-30 июня 2015г. (регуляризация Байеса, архитектура NARX)

Как видно из графиков, они значительно отличаются по форме и величине пиковых нагрузок от графиков, полученных в зимний период времени. Это в большей мере связано с тем, что значительно отличаются

часы, на которые приходится пик нагрузок. В зимние месяцы это вечерние часы, так как заход солнца происходит достаточно рано, повсеместно включается освещение, а также в связи с заходом солнца может значительно снижаться температура.

В летние месяцы пиковые часы нагрузок приходятся на середину дня, так как своего пика достигает температура окружающего воздуха, вследствие чего в помещениях повсеместно начинают включаться системы кондиционирования воздуха.

Подобную зависимость электропотребления от метеофакторов можно объяснить тем, что в случае энергосбытовой компании значительную долю в ее электропотреблении составляют население и приравненные к нему категории потребителей. В процентном соотношении эта доля может достигать 50% и более.

Стоит также сказать, что выбор именно этого месяца для составления прогнозов обусловлен наличием доступа к данным автоматизированной информационно измерительной системы коммерческого учета электроэнергии за данный период времени, а также за аналогичные периоды предыдущих лет.

Что касается графиков, то отклонения за 29 июня 2015г. составили в процентном соотношении к суммарному электропотреблению за сутки составили 5,74% по модулю (4,66% превышение фактического графика над плановым, -1,08% превышение планового графика над фактическим). В абсолютном выражении отклонения составили 145 966 кВт·ч за операционные сутки.

За 30 июня 2015г. отклонения составили 2,61% в процентном отношении по модулю к суммарному электропотреблению за сутки (1,91% превышения фактического графика над плановым, -0,7% превышение планового графика над фактическим) и в абсолютном выражении отклонения составили 56 880 кВт·ч.

Такие результаты можно назвать удовлетворительными, особенно если говорить о показателях 30 июня. Распознавание формы графиков происходит верно, однако существует проблема прогноза на первые сутки двухдневного периода. Вообще стоит сказать, что составление прогноза на двое суток вперед – это вынужденная мера, которая необходима вследствие того, что программному комплексу Matlab для составления прогноза требуется минимум 100 значений временного ряда.

Если рассматривать все графики, представленные в данной работе, то можно заметить, что уменьшение ошибки происходит к концу временного периода, на который составляется прогноз. Это можно объяснить особенностью обучения искусственной нейронной сети. Вычисление весов синаптических функций происходит последовательно, начиная с первых значений временного ряда. Следовательно, к концу временного периода веса синаптических функций принимают оптимальные значения, нежели в начале процесса обучения. Вследствие этого к концу временного периода происходит снижение ошибки прогнозирования.

В целом, данная проблема не является критической, так как составление прогноза происходит на сутки вперед и прогноз на первый день двухдневного периода можно не учитывать, а использовать только прогнозные значения второго дня. Другими словами, первые сутки прогнозного периода используются сетью для обучения и вычисления оптимальных весовых коэффициентов синаптических функций, а вторые – для предоставления результатов прогнозирования.

3.2 Оценка эффективности введения в прогнозную модель возможности учета погодных данных

Как уже говорилось ранее, наличие большого количества потребителей, относимых к категории населения или приравненных к нему, накладывает свой отпечаток на характер суммарного электропотребления энергосбытовой компании. Большое влияние на поведение графика

электропотребления начинают оказывать различные погодные, а также социально значимые факторы.

К метеофакторам, влияющим на электропотребление, можно отнести температурный режим окружающей среды, степень освещенности, наличие или отсутствие осадков, продолжительность светового дня, а также наличие значительного изменения вышеперечисленных погодных условий в течение краткосрочного периода времени.

Наиболее сильное влияние на электропотребление оказывает температурный режим окружающей среды, так как его изменение непосредственно влечет за собой включение или отключение электроприборов, участвующих в создании комфортных условий работы и проживания населения в много квартирных домах. Соответственно, в летний период это кондиционеры в многоквартирных домах и торговых центрах. В зимний период это различные отопительные установки.

Различная степень освещенности также оказывает значительное влияние на электропотребление, так как длительность работы осветительных приборов напрямую зависит от продолжительности светового дня.

Осадки также значительно влияют на график электропотребления, так как от этого зависит, как много жителей города останутся дома в этот день. Форма графиков, например в ясный и дождливый день значительно различается. Разница в них может достигать 50 000 – 60 000 кВт·ч суммарно за сутки.

К социально значимым факторам, которые также оказывают влияние на электропотребление энергосбытовой компании можно отнести статус дня (рабочий, выходной или праздничный день), включение или выключение централизованного отопления в многоквартирных домах жилых районов города, а также возможные коммунальные аварии, непосредственно затрагивающие население.

В связи с вышеперечисленными факторами, целесообразным было включение возможности учета метеофакторов в модели прогнозирования. Из

всех факторов были выбраны наиболее важные, которые наибольшим образом оказывают влияние на электропотребление. Перечислим эти факторы:

- температура окружающего воздуха;
- естественная освещенность;
- статус дня.

Так, например, для описания освещенности были приняты следующие обозначения: 1 – ясно; 2 – 20-30% облачность; 3 – 40-60% облачность; 4 – 70-80% облачность; 5 – 90-100% облачность. Значения температуры вносились в базу данных исходя из фактических значений в каждом получасовом отрезке времени. Статус дня был обозначен следующим образом: 1 – понедельник; 2 – вторник, среда, четверг; 3 – пятница; 4 – суббота; 5 – воскресенье. Разделение дней недели на группы является необходимым, т.к. каждая группа дней имеет особую форму графика электропотребления и не учет этого фактора при прогнозировании может привести к ошибочному прогнозу, а в случае с использованием ИНС – неправильному ее обучению.

3.3 Оценка эффективности прогнозирования выбранной модели нейронной сети с возможностью учета метеофакторов

Для введения возможности учета метеофакторов в модель прогнозирования электропотребления, необходимо было использовать измененную архитектуру искусственной нейронной сети нелинейной авторегрессии. Данная архитектура позволяет производить параллельные вычисления одновременно нескольких параметров, чего и требует поставленная задача. На рисунке 24 представлен общий вид нейронной сети нелинейной авторегрессии с возможностью параллельных вычислений.

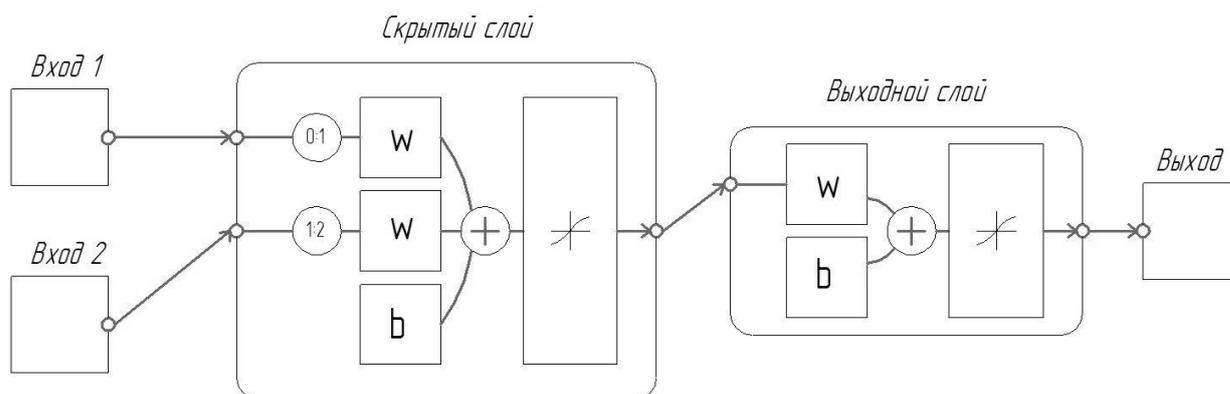


Рисунок 24 – Искусственная нейронная сеть нелинейной авторегрессии (NARX) с возможностью параллельных вычислений

Результатом прогнозирования с помощью данной сети стал график на 29-30 июня 2015г.

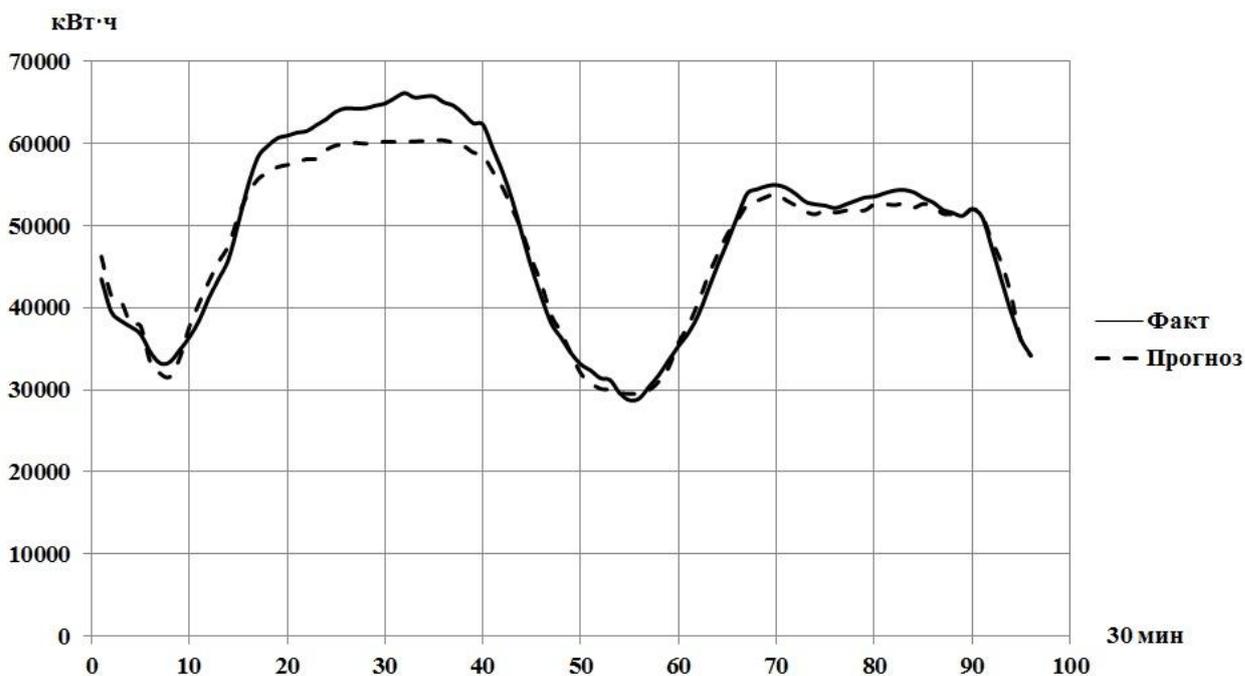


Рисунок 25 - Фактический и прогнозный графики электропотребления энергосбытовой компании для 29-30 июня 2015г. (регуляризация Байеса, архитектура NARX с возможностью параллельных вычислений)

На рисунке 25 представлены прогнозный и фактический графики электропотребления энергосбытовой компании за 29-30 июня 2015г. В

процентном соотношении к суммарному электропотреблению за сутки ошибка для 29 июня составила 5,39% (4,48% превышение факта над планом, 0,91% превышение плана над фактом). В абсолютном выражении ошибка прогнозирования составила 136 966 кВт·ч.

Для 30 июня по тем же параметрам ошибка составила 2,09% суммарно за сутки (1,47% превышение факта над планом, 0,62% превышение плана над фактом). Абсолютное отклонение прогнозного графика электропотребления от фактического суммарно за операционные сутки составило 45 562 кВт·ч.

Полученные данные свидетельствуют о том, что введение возможности учета метеофакторов в модель прогнозирования электропотребления повышает эффективность работы искусственной нейронной сети. Что касается самих результатов, то в общем их можно охарактеризовать как положительные. В особенности прогнозный график на 30 июня 2015г. имеет отклонения от фактического, величина которых является удовлетворяющей требованиям эффективной работы энергосбытовой компании на оптовом рынке электроэнергии. Данный критерий, установленный руководством компании, составляет 5%. Также необходимо сказать о том, что такие показатели точности прогнозирования позволяют данной модели конкурировать с показателями точности, получаемыми в ходе прогнозов экспертами-аналитиками. Априори считается, что данный метод является наиболее точным и эффективным при составлении прогнозов подобного рода. В сравнение же с показателями точности, полученными при использовании других программных комплексов, ошибка прогнозирования в 2,09% - это очень хороший результат, учитывая, что в основной массе автоматизированные программные комплексы дают ошибку в 5-6%.

Вывод по главе 3

В данной главе магистерской диссертации исследована возможность введение в модель прогнозирования на базе искусственной нейронной сети возможности учета метеофакторов, влияющих на электропотребление энергосбытовой компании. Проведен сравнительный анализ результатов прогнозирования с использованием данной модели и модели без возможности учета метеофакторов. Результаты прогнозирования показали, что произошло увеличение точности прогнозирования. Это позволяет сделать вывод о том, что искусственная нейронная сеть построена верно и данная модель прогнозирования является эффективной и конкурентноспособной.

Заключение

В исследовательской работе рассмотрены вопросы работы энергосбытовой компании на оптовом рынке электроэнергии (мощности). В частности рассмотрен вопрос краткосрочного прогнозирования электропотребления для участия в торгах на ОРЭ(М).

Проведен анализ существующих методик прогнозирования, рассмотрены перспективные направления развития данной области научного знания. Анализ показал, что на данный момент наиболее распространенной методикой прогнозирования электропотребления на предприятиях является методика, основанная на использовании знаний и опыта эксперта-аналитика, который в ручном режиме составляет прогноз на сутки вперед, используя при этом данные об электропотреблении за аналогичные периоды времени.

В качестве альтернативы данной методики прогнозирования была рассмотрена перспективная методика прогнозирования на основе использования искусственных нейронных сетей. В работе описана предыстория создания данного метода, приведены теоретические сведения, позволяющие понять принцип работы подобных моделей, области их использования, а также перспективы развития данной области науки.

На следующем этапе работы были решены задачи реализации искусственной нейронной сети на программном уровне. Были рассмотрены различные программные комплексы, позволяющие решать подобные задачи, и из всего многообразия программного обеспечения была выбрана продукция компании MathWorks. Модуль программы Matlab Neural Network Toolbox хорошо себя зарекомендовал и широко используется в научном сообществе для решения подобных задач.

Другой не менее важной задачей стало выбор конфигурации искусственной нейронной сети. Были рассмотрены различные архитектуры сетей и принципы их обучения, использующиеся для решения подобных задач. После тщательного анализа результатов, для прогнозирования была

выбрана нейронная сеть с нелинейной авторегрессией и обучением по принципу регуляризации Байеса.

С учетом специфики работы энергосбытовой компании, а именно большой доли потребителей, относящихся к категории населения и приравненных к нему, необходимым условием эффективной работы модели прогнозирования электропотребления было введение в нее возможности учета погодных факторов. После проведенной работы по адаптации сети для создания возможности учета этих данных, ошибка прогнозирования составила 2,09% отклонений по модулю суммарно за операционные сутки. Это является достаточно хорошим результатом, в сравнении с другими программными комплексами.

Исходя из полученных результатов, можно сделать вывод о том, что созданная модель прогнозирования способна повысить эффективность работы энергосбытовой компании на оптовом рынке электроэнергии (мощности).

Список использованных источников

1. **Постановления Правительства Российской Федерации.** Функционирование розничных рынков электрической энергии, полном и (или) частичном ограничении режима потребления электрической энергии от 04.05.2012 № 442 // Российская газета.
2. **Постановления Правительства Российской Федерации.** Утверждение Правил оптового рынка электрической энергии и мощности и о внесении изменений в некоторые акты Правительства Российской Федерации по вопросам организации функционирования оптового рынка электрической энергии и мощности от 04.12.2010 № 1172 // Парламентская газета.
3. **Российская Федерация. Законы.** Об энергосбережении и о повышении энергетической эффективности и о внесении изменений в отдельные законодательные акты Российской Федерации [Текст] : федер. закон : [принят Гос. Думой 11 ноября 2009 г. : одобр. Советом Федерации 18 ноября 2009 г.]. – М. : - (Актуальный закон).
4. **ГОСТ 32144-2013.** Электрическая энергия. Совместимость технических средств электромагнитная. Нормы качества электрической энергии в системах электроснабжения общего назначения [Текст]. – Введ. 2013-03-25. – М.: Стандартинформ, 2014.
5. Правила устройств электроустановок [Текст] : утв. М-вом энергетики Рос. Федерации 08.07.02 : введ. в действие с 01.01.2003 – 7-е изд. – М.: ЭНАС, 2005.
6. Карпова, Т.С. Базы данных: модели, разработка, реализация [Текст] / Т.С. Карпова ; - СПб.: Питер, 2001. С. 286-289.
7. Боровиков, В.П. Нейронные сети. Statistica Neural Networks. Методология и технологии современного анализа данных [Текст] / В.П. Боровиков. - 2 изд. - М.: Горячая линия - Телеком, 2008. С. 114.
8. Хайкин, С. Нейронные сети: полный курс [Текст] / С. Хайкин - 2 изд. - М.: Вильямс, 2006. С. 89-102.

9. Галушкин, А.И. Нейронные сети: основы теории [Текст] / А.И. Галушкин - М.: Горячая Линия - Телеком, 2010. С. 56.
10. Аксенов, С.В. Организация и использование нейронных сетей (методы и технологии) [Текст] / С.В. Аксенов, В.Б. Новосельцев - Томск: НТЛ, 2006. С.15-19.
11. Бодянский, Е.В. Искусственные нейронные сети: архитектуры, обучение, применение [Текст] / Е.В. Бодянский, О.Г. Руденко - Харьков: Телетех, 2004. С. 304.
12. Круглов, В.В. Нечеткая логика и искусственные нейронные сети [Текст] / В.В. Круглов, М.И. Дли, Р.Ю. Голунов - М.: Издательство Физико-математической литературы, 2001.
13. Толкачев, С.Ф. Нейронное программирование диалоговых систем [Текст] / С.Ф. Толкачев - СПб.: Корона-Век, 2016.
14. Головкин, В.А. Нейронные сети: обучение, организация и применение [Текст] / В.А. Головкин - 4 изд. – М.: ИПРЖР, 2002.
15. Уоссермен Ф. Нейрокомпьютерная техника: Теория и практика [Текст] / Ф. Уоссермен - М.: Издательство Физико-математической литературы, 1992.
16. Hudson Beale, M. Neural Network Toolbox User's Guide [Текст] / М. Hudson Beale, М. Hagan, Н. Demuth - Natick: MathWorks, 2014. p. 22.
17. Hassoun, М.Н. Fundamentals of Artificial Neural Networks [Текст] / М.Н. Hassoun - Massachusetts Institute of Technology: MIT Press, 1995. pp. 254-271.
18. Дерябин, В.В. Нейросетевые алгоритмы в задаче счисления пути судна [Текст] / В.В. Дерябин // Наука и транспорт - 2012. - №2. - С. 74-75.
19. Рузанов, Р.В. Прогнозирование электропотребления энергосбытовой компании с использованием искусственных нейронных сетей [Текст] / Р.В. Рузанов // Энергетика. Проблемы и перспективы развития - Тамбов: ФГБОУ ВПО «ТГТУ» - 2015.

20. Кретов, Д.А. Анализ эффективности использования искусственной нейронной сети для прогнозирования электропотребления энергосбытовой компании [Текст] / Д.А. Кретов, Р.В. Рузанов // Молодые ученые - ускорению научно-технического прогресса в XXI веке - Ижевск: ИННОВА - 2015.

21. Кретов, Д.А. Краткосрочное прогнозирование электропотребления энергосбытовой компании с использованием искусственной нейронной сети [Текст] / Д.А. Кретов, Р.В. Рузанов // Электроэнергетика глазами молодежи - Иваново: ФГБОУ ВПО «Ивановский государственный энергетический университет им. В.И. Ленина» - 2015.

22. Кретов, Д.А. Прогнозирование электропотребления энергосбытовой компании с использованием искусственной нейронной сети [Электронный ресурс] / Д.А. Кретов, Р.В. Рузанов // Инженерный вестник Дона. – 2015. - №2.

23. Галушка, В.В. Формирование обучающей выборки при использовании искусственных нейронных сетей в задачах поиска ошибок баз данных [Электронный ресурс] / В.В. Галушка, В.А. Фатхи // Инженерный вестник Дона. – 2013. - №2.

24. Пучков Е.В. Сравнительный анализ алгоритмов обучения искусственной нейронной сети [Электронный ресурс] / Е.В. Пучков // Инженерный вестник Дона. – 2013. - №4.

25. Администратор торговой системы. –
(<https://www.atsenergo.ru/results/rsv/indexes.html>).

26. Ассоциация НП Совет Рынка. - (<http://www.np-sr.ru/regulation/joining/reglaments/index.html>).

27. StatSoft. –
(http://www.statsoft.ru/products/STATISTICA_Neural_Networks.html).

28. Mathworks. – (<http://www.mathworks.com/products/neural-network.html>).

29. Центр финансовых расчетов. –
(http://cfrenergo.ru/calculations_OREM/calculations.html).

30. Министерство энергетики и жилищно-коммунального хозяйства Самарской области. –
(http://www.minenergo.samregion.ru/norm_base/prikaz_regulirovanae2013/23700.html)