

Направление подготовки 09.03.03 «Прикладная информатика»
Профиль «Прикладная информатика в топливно-энергетическом комплексе»
Методическое обеспечение РПД
Б1.В.ДВ.03.01 «Интеллектуальные информационные системы»



**Филиал федерального государственного бюджетного образовательного учреждения
высшего образования
«Национальный исследовательский университет «МЭИ»
в г. Смоленске**

**МЕТОДИЧЕСКИЕ МАТЕРИАЛЫ ДЛЯ ОБЕСПЕЧЕНИЯ
ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА**

Направление подготовки: **09.03.03 «Прикладная информатика»**

Профиль: **«Прикладная информатика в топливно-энергетическом комплексе»**

Уровень высшего образования: **бакалавриат**

Нормативный срок обучения: **4 года**


Форма обучения: **очная**

Год набора: **2023**

Смоленск


*Направление подготовки 09.03.03 «Прикладная информатика»
Профиль «Прикладная информатика в топливно-энергетическом комплексе»
Методическое обеспечение РПД
Б1.В.ДВ.03.01 «Интеллектуальные информационные системы»*

Методические материалы составил:

канд. техн. наук, доцент кафедры
информационных технологий в экономике и управлении  А.Ю. Пучков

«20» января 2023 г.

Заведующий кафедрой информационных технологий в экономике и управлении:


_____ д-р техн. наук, профессор М.И. Дли
подпись _____ ФИО

«08» февраля 2023 г.

МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ЛАБОРАТОРНЫХ РАБОТ ПО ДИСЦИПЛИНЕ

Лабораторная работа №1. Разработка нечеткого аппроксиматора функциональных зависимостей

Цель занятия: приобрести практические навыки аппроксимации функций в среде MatLAB средствами нечеткой логики, входящими в пакет FuzzyLogicToolbox.

Теоретические сведения

Аппроксимация – научный метод, состоящий в замене одних объектов другими, в каком-то смысле близкими к исходным, но более простыми.

Аппроксимация позволяет исследовать числовые характеристики и качественные свойства объекта, сводя задачу к изучению более простых или более удобных объектов (например, таких, характеристики которых легко вычисляются или свойства которых уже известны). В теории чисел изучаются диофантовы приближения, в частности, приближения иррациональных чисел рациональными. В геометрии рассматриваются аппроксимации кривых ломаными. Некоторые разделы математики в сущности целиком посвящены аппроксимации, например, теория приближения функций, численные методы анализа.

Под нечётким множеством A понимается совокупность упорядоченных пар, составленных из элементов x универсального множества X и соответствующих степеней принадлежности (μ):

$$A = \{(x, \mu_A(x)) \mid x \in X\}$$

где $\mu_A(x)$ – функция принадлежности.

За проведение аппроксимации в MatLAB отвечает пакет прикладных программ FuzzyLogic, обеспечивающий поддержку современных методов нечеткой кластеризации и адаптивных нечетких нейронных сетей. Графические средства пакета позволяют интерактивно отслеживать особенности поведения системы. Данный пакет позволяет:

- определить переменные, нечеткие правила и функций принадлежности;
- осуществлять интерактивный просмотр нечеткого логического вывода;
- использовать современные методы: адаптивный нечеткий вывод с использованием нейронных сетей, нечеткая кластеризация;
- проводить интерактивное динамическое моделирование в Simulink.

Задание к лабораторной работе:

Выполнение заданий 1-2 предполагает оформление отчета в текстовом редакторе Word (TimesNewRoman, 12 п., одинарный), в котором отражаются этапы выполнения заданий и результаты.

Задание 1. Изучить методику аппроксимации функций с помощью FuzzyLogicToolbox.

1.1. Выполнить поочередно аппроксимацию двух функциональных зависимостей. Первая функциональная зависимость – резко изменяющаяся функция, задана в таблице 1.1 дискретными значениями y_i . Соответствующие значения x_i надо рассчитать самостоятельно, взяв из заданного диапазона изменения x восемь равноотстоящих точек.

Таблица 1.1 – Виды функций для аппроксимаций

№	Диапазон изменения x , [a, b]	Значения резко изменяющейся функции								Плавная функция $f(x)$
		y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	
1	(-4, 4)	-8	8	4	-9	6	7	-9	7	$\sin(x)$
2	(0, 7)	4	8	3	9	7	1	5	4	$2*\cos(x)$
3	(-3, 6)	-4	-2	-9	8	8	0	9	-4	$\cos(\sin(x))$
4	(-2,-4)	7	1	-2	8	9	3	5	7	x^2
5	(30, 59)	9	2	4	-7	5	6	1	-9	$2-x^3$
6	(10,26)	0	9	5	4	7	8	-3	4	$1/(x+2)$
7	(-30,3)	4	5	-8	2	9	-7	1	3	$\sin(0.5*x)$
8	(-13,2)	9	9	9	-8	-8	9	9	-8	$x+\sin(x)$
9	(2,3)	1	8	2	8	-9	3	-9	6	$x*\sin(3x)$
10	(8,111)	5	-8	9	-4	5	-6	7	-3	$\cos(0.2*x)$

Вторая функциональная зависимость – плавно изменяющаяся. Для нее надо в тех же, ранее найденных, точках x_i из диапазона изменения x рассчитать значения функции $f(x)$ заданной в виде формулы в таблице 1.1. Провести аппроксимацию полученных точек.

1.2. Провести исследование влияния параметров функций принадлежности на качество аппроксимации. С этой целью для резкоменяющейся функции $y(x_i)$ повторить аппроксимацию при трех различных степенях перекрытия функций принадлежности для переменной x . На основе визуального сравнения сделать вывод, при каком перекрытии аппроксимация более точно соответствует реальному графику, построенному в MatLAB по исходным табличным данным. Повторить эту же процедуру исследования для $f(x_i)$, сохранить результаты, построить графики и сделать выводы.

Задание 2. Построить в MatLAB 3D-график поверхности $z(x,y)$.

2.1. Из таблицы 1.2 при изменении каждого аргумента от -5 до 5 с шагом $h=0.1$. Далее, построить нечеткий аппроксиматор для $z(x,y)$ при изменении каждого аргумента от -5 до 5 с шагом $h=2.5$. После построения аппроксиматора просмотреть выходную поверхность, которую он формирует и визуально сравнить с ранее построенным 3D-графиком $z(x,y)$.

Таблица 1.2. – Вид функции $z(x,y)$ для аппроксимации

№ компьютера	1	2	3	4	5	6
$z(x,y)$	x^2+y^2	$\exp(x^2+y^2)$	$\sin(x^2+y^2)$	$\lg(x^2+y^2 +1)$	$2x^2+7y^3-xy$	$\sin(x)\cos(-y)$
№ компьютера	7	8	9	10	11	12
$Z(x,y)$	$2x^2+5\cos(y^2)$	$\exp(-x^2-y^2)$	$\ln(x^2+2+ y ^3)$	$y^{2*\sin(x)}$	$1/(x^2+ y ^3+2)$	$\sin(x)/(2+\cos(y))$

2.2. Написать программу MatLAB, которая использует созданный аппроксиматор $z(x,y)$ и выполняет следующие действия:

– выводит на экран матрицу значений функции $z(x,y)$ при изменении каждого аргумента от -5 до 5 с шагом $p=2.5$ и матрицу, содержащую результаты аппроксимации $z_a(x,y)$ в тех же точках. Сделать вывод о наличии расхождений в значениях самой функции и ее аппроксимации.

– рассчитывает среднеквадратичную ошибку (СКО) аппроксимации:

$$СКО = \sqrt{\frac{1}{n \cdot m} \sum_{i=1}^n \sum_{j=1}^m (z_{i,j} - z_{a,i,j})^2}$$

где $n \times m$ – число точек, в которых проводится аппроксимация, равное произведению количества строк n и столбцов m матрицы значений функции $z(x,y)$;

$z_{i,j}$ – точные значения функции $z(x_i,y_j)$,

$z_{a,i,j}$ – аппроксимированные значения функции $z(x_i,y_j)$.

– расчет СКО произвести на заданном отрезке $[-5, 5]$ для двух значений шага $h=2,5$ и $h=0,05$. Сравнить значения и сделать выводы;

– построить в одном графическом окне, но в разных подокнах графики $z(x,y)$ и $z_a(x,y)$ (один под другим) в заданном диапазоне при $h=0,05$. Визуально сравнить график $z_a(x,y)$ с выходной поверхностью, выдаваемой в модуле разработки нечетких систем.

Контрольные вопросы

1. Дать определение нечеткого множества.
2. Привести примеры записи нечеткого множества.
3. Основные характеристики нечетких множеств.
4. Дать определение функции принадлежности.
5. Основные возможности пакета FuzzyLogicToolbox.

Лабораторная работа №2. Разработка простой нечеткой экспертной системы

Цель занятия: приобрести практические навыки разработки простой нечеткой экспертной системы (НЭС) в среде MatLAB средствами, входящими в пакет FuzzyLogicToolbox.

Теоретические сведения

В условиях недостатка статистических данных и высокой неопределенности внешней среды методы, основанные на экспертных системах, использующих опыт и знания сотрудников, неточные предварительные данные, предположения, могут являться основой устойчивого экономического развития компании.

Экспертная система (ЭС) – компьютерная система, архитектура которой позволяет автоматизировать достоверные рассуждения человека – эксперта в конкретной предметной области.

ЭС – диалоговая система. Основная цель ЭС – тиражирование и обобщение знаний экспертов, что приводит к удешевлению и ускорению процесса принятия решений, может повысить достоверность результатов за счет минимизации влияния человеческого фактора.

Основные области применения ЭС:

– проблемы в области электротехники, телекоммуникаций, автоматизации управления техническими процессами, промышленным производством; проектирование интегральных схем;

- военные приложения;
- медицина – диагностика заболеваний, оптимизация методов лечения;
- информационные образовательные системы;
- финансовый рынок, рынок недвижимости;
- банковское дело;
- инвестиционный анализ, задачи рационального распределения ресурсов на производстве и т.д.

Нечеткая экспертная система – экспертная система, в которой используется представление знаний в виде лингвистических переменных и нечетких правил, а также применяются алгоритмы нечеткого вывода для получения новых знаний. Можно выделить следующие достоинства нечетких экспертных систем – нелинейность, возможность использовать неточные данные, удобство для получения и обработки экспертных мнений.

Задача системы нечеткого вывода состоит в том, чтобы, руководствуясь базой правил, для фактических (часто точных численных) значений входных переменных определить значения выходных переменных в точной форме или в виде нечеткого множества. Для этого разработан ряд алгоритмов нечеткого вывода, соответствующих специфике поставленной задачи. Фундаментальные принципы нечеткого вывода первым изложил в своих работах английский математик Е. Мамдани – разработанная им схема вывода наиболее интуитивно понятна и позволяет строить адекватные модели, наилучшим образом отражающие суть решаемой задачи. Ключевыми понятиями в алгоритме Мамдани являются термины «фаззификация» и «дефаззификация».

Фаззификация (введение нечеткости) – это установка соответствия между численным значением входной переменной системы нечеткого вывода и значением функции принадлежности соответствующего ей терма лингвистической переменной.

Дефаззификация в системах нечеткого вывода – это процесс перехода от функции принадлежности выходной лингвистической переменной к её четкому (числовому) значению. Цель дефаззификации – используя результаты аккумуляции всех выходных лингвистических переменных, получить количественные значения для каждой выходной переменной.

Задание к лабораторной работе:

Выполнение заданий 1-4 предполагает оформление отчета в текстовом редакторе Word (TimesNewRoman, 12 п., одинарный), в котором отражаются этапы выполнения заданий и результаты (ответы на контрольные вопросы, скриншоты кодов, результатов работы и т.д.).

Задание 1. Изучить методику разработки НЭС в среде MatLAB с помощью FuzzyLogicToolbox.

Задание 2. Выполнить разработку блока нечеткого вывода (БНВ) №1 под именем bank со следующими характеристиками:

Входы системы:

– вход № 1- «проценты» (изменяются от 0% до 30%). Имеет три терма «низкие проценты» (функция принадлежности zmf), «средние проценты» (trimf), «высокие проценты» (smf);

– вход № 2 – «уставной капитал» (изменяется от 0 до 100 мл. руб.). Имеет три терма «малый УК» (zmf), «средний УК» (trapmf), «высокий УК» (smf). Функции принадлежности трапециевидальные;

– вход № 3 – «возраст банка» (изменяется от 0 до 10 лет), имеет четыре термина «малый» (zmf), «средний» (gausmf), «большой» (gausmf), «очень большой» (smf).

Выходы системы (функции принадлежности выбрать самостоятельно):

– выход № 1 – «прибыль» (изменяется от 0 до 1, безразмерная величина). Имеет три термина «малая», «средняя», «высокая»;

– выход № 2 – «риск» (изменяется от 0 до 1, безразмерная величина). Имеет три термина «малый риск», «средний риск», «высокий риск».

Алгоритм вывода – Mamdani. Заполнить базу правил (базу знаний (БЗ)), сформировав не менее 16 правил. В БЗ включить несколько правил, содержащих не все входные переменные. Проверить работоспособность системы.

Задание 3. Выполнить разработку БНВ №2 под именем rezultat со следующими характеристиками:

Входы системы:

– вход № 1- «прибыль» (изменяется от 0 до 1, безразмерная величина). Имеет четыре термина «малая», «средняя», «высокая», «очень высокая»;

– вход № 2 – «риск» (изменяется от 0 до 1, безразмерная величина). Имеет два термина «малый риск», «большой риск»;

Выходы системы:

– выход № 1 – «совет» (изменяется от 0 до 1, безразмерная величина). Имеет два термина «вложение средств целесообразно» (smf), «вложение средств не целесообразно» (zmf);

Алгоритм вывода – Mamdani. Заполнить БЗ не менее чем 16-ю правилами. Проверить работоспособность системы.

Задание 4. Написать программу для НЭС.

4.1. Программа должна объединять ранее созданные БНВ таким образом, что выходы БНВ №1 «bank» будут входами для БНВ №2 «resultat». Сохранить программу и проверить ее работоспособность.

4.2. Изменить программу таким образом, чтобы она показывала значение получаемого риска и, если он превышает 0,5, предупреждать пользователя: «риск большой, хотите ли рискнуть?». Если пользователь рискует, программа продолжает работать и результаты БНВ №1 подаются на БНВ №2. В противном случае работа программы заканчивается. Проверить работоспособность программы.

Контрольные вопросы

1. Дайте определение термину «экспертная система».
2. Выделите основные направления применения экспертных систем.
3. Дайте определение термину «нечеткая экспертная система».
4. Выделите основные достоинства нечетких экспертных систем.
5. В чем заключается цель дефизификации?

Лабораторная работа №3. Разработка гибридной нечеткой экспертной системы

Цель занятия: получение практических навыков построения гибридных нечетких экспертных систем (НЭС), путем создания связи между БД, реализованной в среде Access, и средой MatLAB.

Теоретические сведения

Для подключения к среде MatLABлюбой базы данных необходимо скачать и установить ODBC или JDBC.

ODBC (OpenDatabaseConnectivity) – это программный интерфейс (API) доступа к базам данных, разработанный компанией Microsoft в сотрудничестве с SimbaTechnologies на основе спецификаций CallLevelInterface (CLI). С помощью ODBC прикладные программисты могли разрабатывать приложения для использования одного интерфейса доступа к данным, не беспокоясь о тонкостях взаимодействия с несколькими источниками. Это достигается благодаря тому, что поставщики различных баз данных создают драйверы, реализующие конкретное наполнение стандартных функций из ODBC с учётом особенностей их продукта. JDBC (JavaDatabaseConnectivity) – это платформенно независимый промышленный стандарт взаимодействия Java-приложений с различными СУБД. Данный стандарт основан на концепции драйверов, позволяющих получать соединение с базой данных по специально описанному URL. Драйверы загружаются динамически (во время работы программы) и, загрузившись, драйвер сам регистрирует себя и вызывается автоматически, когда программа требует URL, содержащий протокол, за который драйвер отвечает. Необходимо отметить, что для организации связи базы данных, организованной на базе MSAccess, и среды MatLAB подходит только ODBC.

Для организации связи между средой MatLAB и базой данных в MSAccess для Windows 7, необходимо выполнить следующие действия:

- установить ODBC и запустить его;
 - в появившемся диалоговом окне выбрать MS AccessDatabase, нажать «Добавить»;
 - в появившемся окне выбрать пункт «MicrosoftAccessDriver» и нажать «Добавить»;
 - ввести имя для новой базы данных в поле «DataSourceName», после чего в окне выбора выбрать ее и нажать «Select»;
 - нажать «ОК», после чего проверить установленное соединение и закрыть ODBC.
- Связь между новой базой данных и MatLAB установлена.

Задание к лабораторной работе:

Выполнение заданий 1-4 предполагает оформление отчета в текстовом редакторе Word (TimesNewRoman, 12 п., одинарный), в котором отражаются этапы выполнения заданий и результаты (ответы на контрольные вопросы, скриншоты кодов, результатов работы и т.д.).

Задание 1. На основе теоретических материалов организовать связь между средой MatLAB и MSAccess.

Задание 2. В MSAccess создать две таблицы – «tovar» и «tip_tovara». Структура данных таблиц приведена в таблицах 3.1-3.2. Заполните таблицы, занеся в таблицу «tip_tovara» минимум пять записей, а в таблицу «tovar» минимум по три записи на каждый тип товара.

Таблица 3.1 – Структура таблицы «tovar»

<i>Имя поля</i>	<i>Назначение поля</i>	<i>Какую информацию будет содержать поле</i>
id_tovar	Ключевое поле	
tip_tovar	Наименование типа товара	Тип товара из имеющегося перечня типов.
name_tovar	наименование товара	Конкретная модификация и марка товара в соответствии с типом товара. Тип товара содержится в другой таблице. Например, для типа «телевизор»: «PanasonicTX-21PS70»
opt_cena	оптовая цена	Оптовая цена данного наименования товара, руб
roz_cena	розничная цена	Розничная цена данного наименования товара, руб.

kol_tovar	количество товара	Количество товара на складе фирмы
-----------	-------------------	-----------------------------------

Таблица 3.2 – Структура таблицы «tip_tovara»

Имя поля	Назначение поля	Какую информацию будет содержать поле
id_tip_tovar	Ключевое поле	
name_tupa	Название типа товара	Примеры типов: «телевизор», «холодильник», «мыло» ...
gar_srok	Гарантийный срок	Срок гарантии на данный тип товара в месяцах

Задание 3. Создать программу в среде MatLAB, которая отвечает за выводит список товаров.

Задание 4. Разработать программу НЭС, которая, используя информацию из базы данных, рассчитывает ожидаемую прибыль (ОП) от реализации товаров и рекомендует товар, для которого ОП максимальна.

4.1. Примерная структура фрагмента НЭС, отвечающего за анализ риска и расчет ОП, приведена на рисунке 3.1.

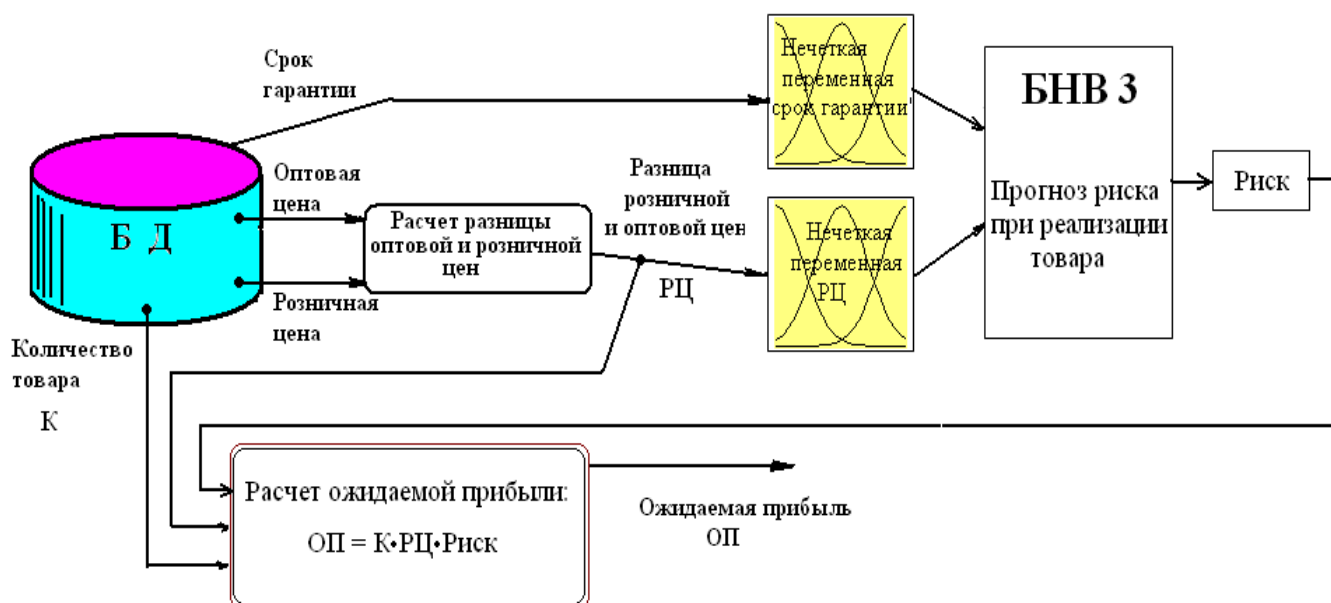


Рисунок 3.1 – Структура фрагмента НЭС, отвечающего за расчет ОП

4.2. Расчет ОП проводить по формуле:

$$ОП = К \cdot РЦ \cdot \text{Риск},$$

где ОП – ожидаемая прибыль,
 К – количество товара на складе,
 РЦ – разность розничной и оптовой цены, руб,
 Риск – нормированный выход БНВ3 «риск».

4.3. Создать базу знаний для БНВ №3, куда включить все возможные сочетания термов входных переменных. Выходом БНВ №3 сделать нормированным от 0 (малый риск) до

1 (большой риск). Предусмотреть, чтобы результатом работы программы НЭС стало не абстрактное число, а символьное предложение, в котором указывается рекомендуемый тип товара, наименование товара и ожидаемая прибыль. Запустить НЭС и оценить правильность ее рекомендации с точки зрения здравого смысла.

Контрольные вопросы

1. Дайте расшифровку термину ODBC. Назначение данного программного обеспечения?
2. Дайте определение термину JDBC. Кратко опишите его назначение и принцип работы.
3. Что применяется для связи MatLAB и MSAccess – ODBC или JDBC?
4. Перечислите основные действия при подключении MSAccess к MatLAB.
5. Назовите операторы и функции языка структурированных запросов SQL и поясните их назначение.

Лабораторная работа №4. Решение задач аппроксимации

Цель занятия: получение практических навыков решения задач аппроксимации в среде MatLAB.

Теоретические сведения

Аппроксимация (или приближение) – научный метод, состоящий в замене одних объектов другими, в каком-то смысле близкими к исходным, но более простыми. Аппроксимация позволяет исследовать числовые характеристики и качественные свойства объекта, сводя задачу к изучению более простых или более удобных объектов (например, таких, характеристики которых легко вычисляются или свойства которых уже известны). В теории чисел изучаются диофантовы приближения, в частности, приближения иррациональных чисел рациональными. В геометрии рассматриваются аппроксимации кривых ломаными. Некоторые разделы математики в сущности целиком посвящены аппроксимации, например, теория приближения функций, численные методы анализа.

Наиболее распространенным методом подбора параметров уравнений является метод наименьших квадратов. Метод наименьших квадратов (МНК) – математический метод, применяемый для решения различных задач, основанный на минимизации суммы квадратов отклонений некоторых функций от искоемых переменных. Суть данного метода заключается в том, что если система уравнений имеет решение, то наименьшее значение суммы квадратов будет равно нулю, и могут быть найдены точные решения системы уравнений аналитически или, например, различными численными методами оптимизации. Если система переопределена, то есть, говоря нестрого, количество независимых уравнений больше количества искоемых переменных, то система не имеет точного решения и метод наименьших квадратов позволяет найти некоторый «оптимальный» вектор x в смысле максимальной близости векторов y и $f(x)$ или максимальной близости вектора отклонений e к нулю.

Некоторые примеры задач аппроксимации и возможные варианты их решения:

1. Для приближённого вычисления интеграла используется формула прямоугольников или формула трапеций, или более сложная формула Симпсона. Фактически при этом происходит приближение подинтегральной функции ступенчатой функцией или вписанной ломаной, интеграл от которой считается мгновенно.

2. Для вычисления значений сложных функций часто используется вычисление значения отрезка ряда, аппроксимирующего функцию.

3. Для обработки экспериментальных или натуральных данных: аппроксимирующая функция ограничена диапазоном заданных точек и служит в качестве только интерполирующей зависимости; аппроксимирующая функция выступает в роли физического закона и с её помощью допускается экстраполировать переменные.

Аппроксимация Паде – классический метод рациональной аппроксимации аналитических функций, названный по имени французского математика Анри Паде. Метод заключается в представлении функции в виде отношения двух полиномов, коэффициенты которых определяются коэффициентами разложения функции в ряд Тейлора.

Задание к лабораторной работе:

Выполнение заданий 1-2 предполагает оформление отчета в текстовом редакторе Word (TimesNewRoman, 12 п., одинарный), в котором отражаются этапы выполнения заданий и результаты (ответы на контрольные вопросы, скриншоты кодов, результатов работы и т.д.).

Задание 1. Изучите функцию $y = \text{tg}(x)$, заданную на интервале $a < x_0 < b$, непрерывно-дифференцируемую n раз в окрестности точки $x_0 \in (a, b)$. Диапазон изменения значения x взять из таблицы 4.1 согласно номеру занимаемого компьютера.

Таблица 4.1 – Виды функций для аппроксимаций

№	Диапазон изменения x , [a, b]	$f(x)$
1	(-4, 4)	$\sin(x)$
2	(0, 7)	$2 \cdot \cos(x)$
3	(-3, 6)	$\cos(\sin(x))$
4	(-2, -4)	x^2
5	(30, 59)	$2 - x^3$
6	(10, 26)	$1/(x+2)$
7	(-30, 3)	$\sin(0.5 \cdot x)$
8	(-13, 2)	$x + \sin(x)$
9	(2, 3)	$x \cdot \sin(3x)$
10	(8, 111)	$\cos(0.2 \cdot x)$

1.1. Приближая значение функции в окрестности точки x_0 многочленом Тейлора с применением соответствующей аппроксимацией Паде.

Задание 2. Для функции $y = f(x)$, заданной в явном виде на интервале $a < x_0 < b$, построить в окрестности указанной точки многочлен Тейлора и соответствующую аппроксимацию Паде. Отобразите график функции серым жирным пунктиром, ее полином Тейлора красной толстой линией и ее аппроксимацию Паде синей линией. Диапазон изменения значения x , а также вид функции $f(x)$ x взять из таблицы 1.1 согласно номеру занимаемого компьютера.

Контрольные вопросы

1. Дайте определение термину «аппроксимация».
2. Какой метод чаще всего применяется для подбора параметров уравнений?
3. В чем суть метода наименьших квадратов?
4. Назовите наиболее распространенные методы аппроксимации.

5. Приведите примеры задач аппроксимации и способы их решения.

Лабораторная работа №5. Прогнозирование данных с помощью ИНС

Цель занятия: ознакомиться с возможностями ИНС, а также получить практические навыки прогнозирования данных с помощью встроенных инструментов.

Теоретические сведения

Искусственные нейронные сети (ИНС, artificial neural networks, ANN) – математическая модель, а также ее программное или аппаратное воплощение, построенная по принципу организации и функционирования биологических нейронных сетей – сетей нервных клеток живого организма. ИНС состоит из искусственных нейронов (artificial neuron), каждый из которых представляет собой упрощенную модель биологического нейрона. Все, что делает искусственный нейрон – это принимает сигналы со многих входов, обрабатывает их единым образом и передает результат на многие другие искусственные нейроны, то есть, делает то же самое, что и нейрон биологический. В синапсах происходит усиление или ослабление электрохимического сигнала. Связи между искусственными нейронами называются синаптическими, или просто синапсами. У синапса имеется один параметр – весовой коэффициент, в зависимости от его значения происходит то или иное изменение информации, когда она передается от одного нейрона к другому. Именно благодаря этому входная информация обрабатывается и превращается в результат, а обучение нейронной сети основано на экспериментальном подборе такого весового коэффициента для каждого синапса, который и приводит к получению требуемого результата. Структура простейшей нейронной сети представлена на рисунке 6.1. Зелёным цветом обозначены нейроны входного слоя, голубым – нейроны скрытого слоя, жёлтым – нейрон(ы) выходного слоя.

Нейроны входного слоя получают данные извне (например, от сенсоров системы распознавания лиц) и после их обработки передают сигналы через синапсы нейронам следующего слоя.

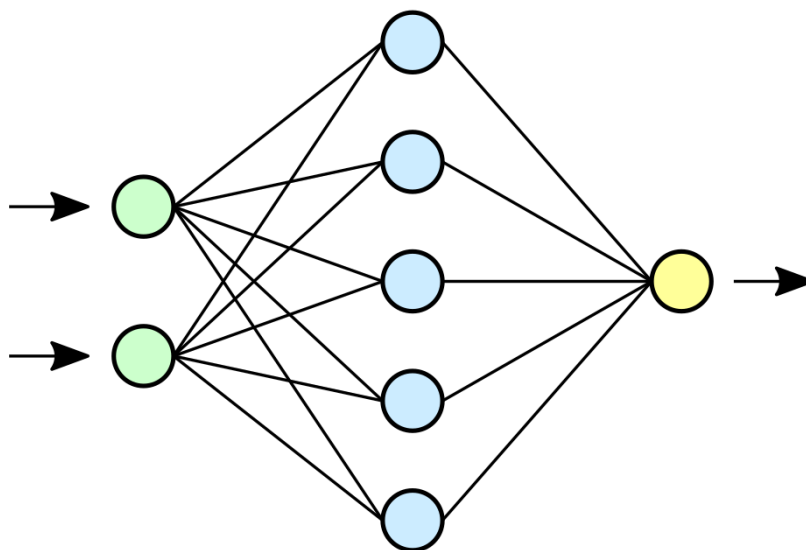


Рисунок 6.1 – Структура ИНС

Нейроны второго слоя (его называют скрытым, потому что он напрямую не связан ни со входом, ни с выходом ИНС) обрабатывают полученные сигналы и передают их нейронам выходного слоя. Поскольку речь идет об имитации нейронов, то каждый процес-

сор входного уровня связан с несколькими процессорами скрытого уровня, каждый из которых, в свою очередь, связан с несколькими процессорами уровня выходного. Такая, простейшая ИНС способна к обучению и может находить простые взаимосвязи в данных. ИНС, способная находить не только простые взаимосвязи, но и взаимосвязи между взаимосвязями имеет намного более сложную структуру.

Можно выделить следующие достоинства ИНС:

- решение задач в условиях неопределенности;
- устойчивость к шумам во входных данных;
- гибкость структуры нейронных сетей;
- высокое быстродействие;
- адаптация к изменениям окружающей среды;
- отказоустойчивость нейронных сетей.

Недостатки ИНС:

- ответ, выдаваемый ИНС, всегда приближительный;
- неспособность принятия решений в несколько этапов;
- неспособность решать вычислительные задачи;
- трудоемкость и длительность обучения.

Задание к лабораторной работе:

Выполнение заданий 1-3 предполагает оформление отчета в текстовом редакторе Word (TimesNewRoman, 12 п., одинарный), в котором отражаются этапы выполнения заданий и результаты (ответы на контрольные вопросы, скриншоты кодов, результатов работы и т.д.).

Задание 1. Изучить структуру ИНС, предоставленную преподавателем.

Задание 2. База данных содержит факты вида: отдыхает (имя, город), Белоруссия(город), Россия(город), женщина (имя), мужчина (имя).

2.1. Для студентов с четным номером в журнале – вывести список женщин, отдыхающих в России.

2.2. Для студентов с нечетным номером в журнале – вывести список мужчин, отдыхающих в Белоруссии.

Задание 3. Разработать ИНС, которая прогнозирует ВВП на 2011 год для выбранной страны. Страну выбрать из таблицы 6.1 в соответствии с номером компьютера.

Таблица 5.2 – Выбор страны для прогноза ВВП

№ компьютера	1	2	3	4	5	6
Страна	Россия	Китай	СССР	Япония	Германия	Великобритания
	7	8	9	10	11	12
	Франция	Китай	Россия	Япония	Германия	Великобритания

3.1. Построить графики прогнозируемых данных и выхода сети по годам, визуально оценив их совпадение с графиками, которые приведены на рисунке 6.2.

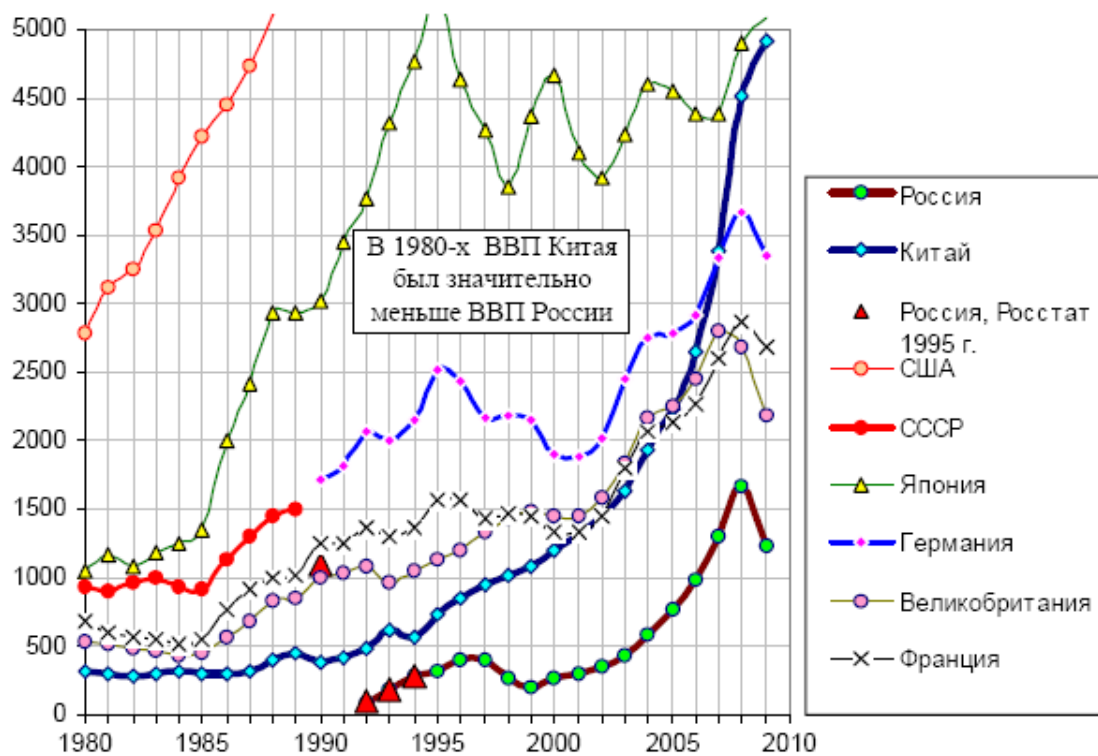


Рисунок 6.2 – ВВП крупнейших стран

3.2. Сравнить спрогнозированные данные с фактическими данными 2011 года.

3.3. Для студентов с нечетным номером в журнале – вывести список книг, изданных в издательстве «Питер» не ранее 2016 года.

Контрольные вопросы

1. Дайте определение термину «искусственная нейронная сеть».
2. Опишите схему простой ИНС.
3. Что такое синапс?
4. Перечислите достоинства ИНС.
5. Перечислите недостатки ИНС.

Лабораторная работа №6. Разработка гибридной нейронной сети для заданной предметной области – часть 1

Цель занятия: закрепить теоретические знания в области искусственных нейронных сетей и получить навыки их практического применения для экономических приложений в среде NNTool и MatLAB.

Теоретические сведения

Типы нейронных сетей, применяемых в NNTool приведены в таблице 6.1.

Таблица 6.1 – Типы нейронных сетей

№ типа	Тип сети	Название сети	Число слоев	Обучаемые параметры
1	Competitive	Конкурирующая сеть	1	$1 \text{ } IW\{1, 1\}, b\{1\}$
2	Cascade-forward back-	Каскадная сеть с прямым распространением сигнала и обратным рас-	2	$IW\{1, 1\}, b\{1\}$ $LW\{2, 1\} \text{ } IW\{2,$

	prop	пространением ошибки		$1\}, b\{2\}$
3	Elman backprop	Сеть Элмана с обратным распространением ошибки	2	$IW\{1, 1\}, b\{1\}$ $LW\{2, 1\}, b\{2\}$ $LW\{2, 1\},$
4	Feed-forward backprop	Сеть с прямым распространением сигнала и обратным распространением ошибки	2	$IW\{1, 1\}, b\{1\}$ $LW\{2, 1\}, b\{2\}$
5	Time delay backprop	Сеть с запаздыванием и обратным распространением ошибки	2	$IW\{1, 1\}, b\{1\},$ $LW\{2, 1\}, b\{2\}$
6	Generalized regression	Обобщенная регрессионная сеть	2	$IW\{1, 1\}, b\{1\},$ $LW\{2, 1\}$
7	Hopfield	Сеть Хопфилда	1	$LW\{1, 1\}, b\{1\}$
8	Linear layer (design)	Линейный слой (создание)	1	$IW\{1, 1\}, b\{1\}$
9	Linear layer (train)	Линейный слой (обучение)	1	$IW\{1, 1\}, b\{1\}$
10	LVQ	Сеть для классификации входных векторов	2	$IW\{1, 1\}, LW\{2, 1\}$
11	Perceptron	Персептрон	1	$IW\{1, 1\}, b\{1\}$
12	Probabalistic	Вероятностная сеть	2	$IW\{1, 1\}, b\{1\},$ $LW\{2, 1\}$
13	Radial basis (exact fit)	Радиальная базисная сеть с нулевой ошибкой	2	$IW\{1, 1\}, b\{1\},$ $LW\{2, 1\}$
14	Radial basis (fewer neurons)	Радиальная базисная сеть с минимальным числом нейронов	2	$IW\{1, 1\}, b\{1\},$ $LW\{2, 1\}, b\{2\}$
15	Self-organizing map	Самоорганизующаяся карта Кохонена	1	$IW\{1, 1\}$

При передаче данных из NNTool в MatLAB не следует использовать команду clear, так как она стирает все данные. Для контроля активных переменных и их типа следует использовать команды whos или who.

Для загрузки выборок в NNTool для обучения и применения сети надо выполнить следующие этапы:

1. Создать файл данных в программе «блокнот» содержащий числа, расположенные в строках и столбцах (как матрица). В строке числа разделяются одним пробелом. Во всех строках должно быть одинаковое количество чисел. Условимся, что значения входов сети, соответствующих одному примеру, располагаем в столбец. Тогда последняя строка в файле должна содержать требуемые значения выхода для соответствующего примера. Таким образом, один пример – это один столбец. Количество обучающих примеров для сети равно числу столбцов данных. Файлу присвоить расширение dat и разместить его в текущей папке MatLAB. По умолчанию это папка Work. Пусть файлу присвоили имя viborka.dat.

2. Читать подготовленный файл в рабочее пространство MatLAB (это там, где командная строка >>), выполнив в командной строке команду $A = \text{load}('viborka.dat')$. В результате в рабочем пространстве MatLAB сформируется матрица A.

3. Выделим из матрицы A входные данные data1 (это все строки, кроме последней) и выходные данные data2 (это последняя строка):

```
r=size(A) % определили размерность двумерного массива A
n=r(1,1) % n=количество строк
```

```
data1=A(1:(n-1),:);
data2= A(n,:);
```

4. Находясь в графической среде NNTool нажать кнопку Import, отметить пункт Importfrom MATLAB Workspace. В результате в соседнем окне отобразится список доступных переменных и матриц в рабочей области (SelectVariable). Выделить мышкой матрицу data1 и в соседнем окне Destination отметить пункт Input. Аналогично загрузить целевые данные data2, отметив пункт Targets (цели).

5. При необходимости сохранить переменные из рабочего пространства MatLAB для будущего использования надо выполнить команду save и присвоить файлу с сохраняемыми данным какое-либо имя, например save apr1_12_2011. В результате в папке Work появится файл apr1_12_2011.mat. Его теперь можно сразу загружать в среду NNTool отметив пункт Loadfromdiskfile. Дальнейшая процедура загрузки аналогична вышеописанной.

6. Передача данных из NNTool в рабочую область MatLAB производится с помощью кнопки Export. Чаще всего бывает надо передать ответы сети на входные данные. Эти ответы заносятся в векторную переменную network1_outputs, которую и надо передать в рабочую область MatLAB используя кнопку Export среды NNTool. После этого, набрав в рабочей области MatLAB имя переменной network1_outputs можно будет увидеть ее значения. Для удобства можно присвоить значение выхода сети переменной с более коротким именем, например: >> Nout= network1_outputs и дальше работать уже с ней.

Имейте в виду, что если запускается на выполнение подготовленная программа, то для того, чтобы ей была доступна для обработки и анализа переменная network1_outputs (или, что тоже самое, Nout в нашем примере), программа не должна содержать в начале команду clear, иначе все переменные обнулятся.

Задание к лабораторной работе:

Выполнение заданий 1-3 предполагает оформление отчета в текстовом редакторе Word (Times New Roman, 12 п., одинарный), в котором отражаются этапы выполнения заданий и результаты (ответы на контрольные вопросы, скриншоты кодов, результатов работы и т.д.).

Задание 1. Построить нейросетевой аппроксиматор для функции $z(x,y)$, вид которой выбрать из таблицы 2.2 в соответствии с номером компьютера. Параметры сети выбрать самостоятельно. Точки аппроксимации взять при изменении каждого аргумента от -5 до 5 с шагом $h=0.25$. Предусмотреть создание обучающей и контрольной выборок. Построить графики истинной функции $z(x,y)$ и ее аппроксимации с помощью нейросети. Рассчитать СКО между истинной функцией и ее аппроксимацией. Графики и СКО занести в отчет.

Задание 2. Выборка статистических данные по России представлены в таблице 5.1. На основании данных таблицы 5.1 необходимо сделать прогноз ВВП на 2010 год с помощью ИНС. Для обучения ИНС взять те поля из таблицы 7.1, которые указаны в таблице 7.2 в соответствии с номером N компьютера. В связи с небольшим объемом исходных данных придется ограничиться только одной, обучающей, выборкой, исключив контрольную.

Таблица 7.1 – Статистические данные

Год	Численность населения, тыс. чел	Кол. Безработных, млн. чел.	Среднемес. доход, тыс. руб	Численность аспирантов в Смоленской области, чел	ВВП, млрд. руб.
1991	148543	–	–	–	–

1992	148704	–	–	46	–
1993	148673	–	45,2	45	–
1994	148366	5,5	206,3	46	–
1995	148306	6,2	515,5	49	–
1996	147976	7	770	54	–
1997	147502	7,8	942,1	81	–
1998	147105	8,5	1013	103	–
1999	146693	9,3	1664	152	4823
2000	145925	7,6	2290,1	186	7305
2001	146304	6,4	3078,4	202	8943
2002	145649	5,7	3947,2	244	10819
2003	144964	6,1	5170,4	281	13208
2004	144168	6	6410,4	321	170027
2005	143474	5,6	8111,9	373	21609
2006	142754	5,3	10196	395	26917
2007	142200	4,6	12602,7	–	33247
2008	141980	4,8	14940,6	–	41276
2009	141900	6,3	16856,9	–	38786

Таблица 7.2 – Списки используемых полей

№ компьютера	1	2	3	4	5	6
Поля	год, среднемес. доход, ВВП	Численность населения, ВВП	Кол. Безработных, ВВП	Численность населения, среднемес. доход, ВВП	Численность аспиранто, ВВП	Год, Численность аспирантов, ВВП
№ компьютера	7	8	9	10	11	12
Поля	Все поля	Численность аспирантов, ВВП	Год, кол. безработных, ВВП	Среднемес. доход, ВВП	год, кол. безработных ВВП	Численность населения, Численность аспирантов, ВВП

Задание 3. Сделать прогноз ВВП на 2010 год по методу экспоненциального сглаживания. Сравнить полученные прогнозы с реальным ВВП за 2010 год, который составил 44939 млрд. руб. Сделать вывод о том, какой прогноз оказался точнее и целесообразно ли применять ИНС – может достаточно ограничиться простым прогнозом по методу экспоненциального сглаживания?

Контрольные вопросы

1. Сколько типов ИНС можно использовать в среде NNTool?
2. Как можно просмотреть графическое изображение созданной ИНС и что значат числа, отображаемые в этом окне?
3. Сколько выходов может иметь ИНС, созданная в NNTool?
4. Где задается количество слоев и число нейронов в слое?
5. Как воспользоваться обученной сетью?

Лабораторная работа №7. Разработка гибридной нейронной сети для заданной предметной области – часть 2

Цель занятия: закрепить теоретические знания в области искусственных нейронных сетей и разработать гибридную нейронную сеть для заданной предметной области.

Теоретические сведения

Гибридная нейронная сеть – это нейронная сеть с четкими сигналами, весами и активационной функцией, но с объединением x_i и ω_i , p_1 и p_2 с использованием t-нормы, t-конормы или некоторых других непрерывных операций. Входы, выходы и веса гибридной нейронной сети – вещественные числа, принадлежащие отрезку $[0, 1]$.

Наиболее значимыми и важными классами нейронных сетей являются:

– сети прямого распространения – все связи направлены строго от входных нейронов к выходным. Такие сети еще называют многослойным персептроном, по аналогии с обычным персептроном Розенблатта, в котором только один скрытый слой;

– рекуррентные нейронные сети или сети обратного распространения – сигнал в таких сетях с выходных нейронов или нейронов скрытого слоя частично передается обратно на входы нейронов входного слоя;

– сеть радиальных базисных функций (RBF) – вид многослойной нейронной сети, имеющий скрытый слой из радиальных элементов и выходной слой из линейных элементов.

– самоорганизующиеся карты или сеть Кохонена. Такой класс многослойных нейронных сетей, как правило, обучается без учителя и успешно применяется в задачах распознавания

– глубокие нейронные сети – сети, состоящие из большого количества слоев, миллионов нейронов и реализующие методы DeepLearning (глубокого обучения). Представляют собой третье поколение нейронных сетей. К ним относятся сверточные, капсульные нейросети, автоэнкодеры, глубинные машины Больцмана. Такие сети успешно применяются для распознавания изображений.

Задание к лабораторной работе:

Выполнение заданий 1-2 предполагает оформление отчета в текстовом редакторе Word (TimesNewRoman, 12 п., одинарный), в котором отражаются этапы выполнения заданий и результаты (ответы на контрольные вопросы, скриншоты кодов, результатов работы и т.д.).

Задание 1. Разработать структуру вопросов (блок-схему) НЭС для предметной области, выданной преподавателем в соответствии с вариантом.

1.1. Начальная оценка работы будет зависеть от сложности НЭС:

– «удовлетворительно» – достаточно реализовать НЭС, рассмотренную на лекции;
– «хорошо» – НЭС должна содержать 6 блоков нечеткого вывода (БНВ) и не менее 5 дополнительных вопросов (их НЭС задает в процессе выполнения программы). Вопросы должны подразумевать как количественные ответы, например, «доходность по акциям составляет 12%», так и качественные, например, «большой убыток», «очень низкое качество» и т.д.;

– «отлично» - НЭС должна содержать не менее 7 блоков БНВ и не менее чем 5 дополнительных вопросов. Вопросы должны подразумевать как количественные ответы, так и качественные.

Задание 2. Разработать программу в среде MatLAB для НЭС, структура вопросов которой была подготовлена в предыдущем задании.

Контрольные вопросы

1. Дайте определение термину «нечеткая экспертная система».
2. Дайте определение термину «гибридная нейронная сеть».
3. Дать графическую интерпретацию алгоритмов нечетких выводов.
4. Целесообразность применения систем нечеткой логики.
5. Описать последовательность действий алгоритмов нечеткого вывода.

Рекомендуемая литература

1 Богданов Е. П. Интеллектуальный анализ данных : учебное пособие. Волгоград : Волгоградский ГАУ : Лань , 2019. 112 с. [электронный ресурс]: <https://e.lanbook.com/book/139228>

2 Федотов А. В. Компьютерное управление в производственных системах: учебное пособие для ВПО : учебное пособие. Санкт-Петербург : Лань, 2020. 620 с. [электронный ресурс]: <https://e.lanbook.com/reader/book/140775/#2>

3 Остроух А. В. Интеллектуальные информационные системы и технологии : монография / А. В. Остроух, А. Б. Николаев. Санкт-Петербург : Лань, 2019. 308 с. [электронный ресурс]: <https://e.lanbook.com/book/115518>

4 Макшанов А.В. Технологии интеллектуального анализа данных : учебное пособие. Санкт-Петербург : Лань, 2019. 212 с. [электронный ресурс]: <https://e.lanbook.com/reader/book/120063/#2>

5 Пенькова Т. Г. Модели и методы искусственного интеллекта : учебное пособие. Красноярск : СФУ, 2019. 116 с. Лань [электронный ресурс]: <https://e.lanbook.com/book/>