

Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Саратовский государственный технический университет
имени Гагарина Ю.А.»

Кафедра «Автоматизация, управление, мехатроника»

РАБОЧАЯ ПРОГРАММА

по дисциплине

Б.1.3.3.2 «Анализ больших данных»

направления подготовки

15.03.04 - Автоматизация технологических процессов и производств

Профиль «Интеллектуальные информационно-управляющие системы»

(для дисциплин, реализуемых в рамках профиля)

форма обучения – очная

курс – 1

семестр – 2

зачетных единиц – 5

часов в неделю – 4

академических часов – 180

в том числе:

лекции – 28

коллоквиумы – 8

практические занятия – 36

лабораторные занятия – не предусмотрены

самостоятельная работа – 108

зачет – не предусмотрен

экзамен – 2-й семестр

РГР – не предусмотрена

курсовая работа – не предусмотрена

курсовой проект – не предусмотрен

1. Цели и задачи дисциплины

Цель преподавания дисциплины: получение теоретической подготовки студентов по современным подходам к анализу больших массивов данных (Big Data) и построения на их основе автоматизированных интеллектуальных информационных систем управления для контроля технологических объектов.

Задачи изучения дисциплины:

1) ознакомление студентов с подходами обработки больших объёмов данных по сравнению со «стандартными» сценариями распознавания образов;

2) на примере современных информационно-управляющих систем усвоить принципы обработки быстро поступающих данных в увеличивающихся объёмах, а также параллельной обработки массивов плохо структурированных данных для одновременного поиска различных закономерностей и особенностей;

3) освоение современных представлений о принципах построения систем обработки больших данных и информационных систем на их основе, их структуре и основах аппаратного и программного обеспечения.

2. Место дисциплины в структуре ООП ВО

Дисциплина Б.1.3.3.2 «Анализ больших данных» входит в часть дисциплин по выбору базового цикла Б.1.3 ФГОС ВО по направлению подготовки бакалавров 15.03.04. В процессе её изучения используются базовые знания студентов по дисциплинам «Физика», «Математика», «Информационные технологии». Знания, полученные при изучении дисциплины «Анализ больших данных» необходимы для изучения дисциплин «Теория динамических систем и сложных сетей в инженерных задачах», «Теория идентификации» и подготовки выпускной квалификационной работы.

3. Требования к результатам освоения дисциплины

Изучение дисциплины направлено на формирование следующих компетенций:

(ПК-1) способностью собирать и анализировать исходные информационные данные для проектирования технологических процессов изготовления продукции, средств и систем автоматизации, контроля, технологического оснащения, диагностики, испытаний, управления процессами, жизненным циклом продукции и ее качеством; участвовать в работах по расчету и проектированию процессов изготовления продукции и указанных средств и систем с использованием современных информационных технологий, методов и средств проектирования:

Знает: основные проблемы и особенности отрасли обработки больших данных; прикладные программные средства для анализа больших данных и построения на их основе систем распознавания образов; области применения методик обработки больших данных;

Умеет: использовать на практике методы и способы анализа больших данных; реализовывать системы многофакторного анализа больших данных; создавать системы адаптивной обработки с автоматизированным поиском критериев для адаптивного анализа больших данных;

Владеет: современными достижениями физико-математических методов и компьютерных технологий по обработке больших данных в целях создания автоматизированных систем управления и слежения в технологических, научных и инженерных задачах;

(ПК-20) способностью проводить эксперименты по заданным методикам с обработкой и анализом их результатов, составлять описания выполненных исследований и подготавливать данные для разработки научных обзоров и публикаций;

Знает: терминологию, цель применения и особенности основных методов оцифровывания, повышения соотношения "сигнал/шум" и обработки больших данных;

Умеет: ставить задачи в обработке больших данных в реальном времени; решать задачи разработки интеллектуальных информационных систем для обработки больших данных; составлять научно-инженерный отчёт по результатам обработки больших данных;

Владеет: способностью оценки выбора возможностей многофакторного анализа больших данных; навыками проведения анализа больших данных; современным научно-инженерным языком для описания результатов проведённого анализа и обработки данных для подготовки научных публикаций.

4. Распределение трудоемкости (час.) дисциплины по темам и видам занятий

№ модуля	№ недели	№ темы	Наименование темы	Часы/ Из них в интерактивной форме				
				Всего	Лекц.	Коллоқ.	Пр.з.	СРС
1	2	3	4	5	6	7	8	9
1	1-5	1	Понятие "большие данные" (Big Data).	40	6	2	4	28
2	6-11	2	Методики анализа больших данных.	52	10	2	10	30
3	12-15	3	Современные программные реализации анализа больших данных.	58	8	2	18	30
4	16-18	4	Когнитивные вычисления в анализе больших данных.	30	4	2	4	20
			Итого	180 /36	28	8	36/ 36	108

5. Содержание лекционного курса

№ темы	Всего часов	№ лекций	Тема лекции. Вопросы, обрабатываемые на лекции	Учебно-методическое обеспечение
1	2	3	4	
1	2	1	Понятие "большие данные" (Big Data). Краткая историческая справка появления проблемы обработки больших данных.	1, 2,16
1	4	2-3	Большие данные: области использования. Обзор по современным применениям в научно-инженерных задачах. Принципы адаптации алгоритмов анализа больших данных. Оценка перспектив развития теории и прикладных методик для анализа больших данных в различных областях естественно-научного знания.	1,-4, 6, 16
2	4	4-5	Методики анализа больших данных, применяемые как в больших, так и в меньших по объёму массивов данных. А/В-тестирование. Регрессионный анализ. Association rule learning. Cluster analysis. Classification. Crowdsourcing. Data mining.	5,8, 16
2	6	6-8	Анализ больших данных на базе искусственных нейронных сетей (ИНС). Принципы построения ИНС. Классификация ИНС. Различные методы обучения ИНС. Подходы к анализу больших данных на базе генетических алгоритмов (Genetic algorithms) и самообучения ИНС (Machine learning).	2,4,16
3	4	9-10	Принципы построения информационных систем и систем распознавания образов для анализа больших данных. Выбор параметров информационных систем с учетом производственных задач.	3,4,5,16
3	4	11-12	Структура современной информационной системы автоматизированного технологического оборудования. Визуализация результатов анализа больших данных и их дальнейшая постобработка.	3, 16
4	4	13-14	Развитие ИНС-подходов в рамках создания когнитивных моделей анализа больших данных. Идея искусственного интеллекта. Облачные технологии. Параллельные вычисления.	2,4,16

6. Содержание коллоквиумов

№ темы	Всего часов	№ коллоквиума	Тема коллоквиума. Вопросы, обрабатываемые на коллоквиуме	Учебно-методическое обеспечение
1	2	3	4	5
1	2	1	Понятие большие данные. Основные критерии дифференциации больших и "обычных" данных.	1-5,7
2	2	2	А/В-тестирование. Регрессионный анализ. Особенности применения в больших данных.	1-5,7
3	2	3	ИНС. Методы обучения для сетей различной архитектуры.	1-5,7

4	2	4	ИНС на базе облачных технологий: проблемы самообучения.	1-5
---	---	---	---	-----

7. Перечень практических занятий

№ темы	Всего часов	№ занятия	Тема практического занятия. Задания, вопросы, отрабатываемые на практическом занятии	Учебно-методическое обеспечение
1	2	3	4	5
1,2	6	1-3	Дифференциация больших и малых массивов данных. Основные методы традиционной обработки данных. Проблемы цифрового анализа больших данных на базе традиционных подходов.	1,3-5,8,16
1,2	8	4-7	Современные методики обработки больших данных. Реализация одной из методик в режиме постобработки больших данных. Написание отчёта по результатам проделанной работы. Освоение научно-инженерной терминологии, используемой в данной отрасли.	1,3-5,8,16
3	8	8-11	Программирование ИНС в среде MatLab. Создание сетей различной архитектуры.	8,9,16
3	4	12-13	Реализация различных обучающих алгоритмов для ИНС с варьирующейся архитектурой. Разработка адаптации архитектуры ИНС к изменению задач обучения.	6,8,9,16
3,4	10	14-18	Создание системы самообучения ИНС. Выбор оптимальной архитектуры ИНС для данных целей. Обработка экспериментальных больших данных в целях распознавания состояния сложной системы на базе созданной ИНС. Написание обзорной работы по результатам проведенных работ по исследованию возможностей когнитивного анализа данных на базе созданного ИНС.	8,9,16

8. Перечень лабораторных работ

Не предусмотрены

9. Задания для самостоятельной работы студентов.

№ темы	Всего Часов	Задания, вопросы для самостоятельного изучения (задания)	Учебно-методическое обеспечение
1	2	3	4
1	28	Обработка больших данных в промышленности. Управление процессами распознавания объектов. Адаптивные системы обработки данных. Эффективность систем распознавания.	1,2,4,10-15
2	30	Системы распознавания объектов в технологических устройствах на основе методов анализа больших данных	3,4,9-15
3	30	Исторический обзор создания и развития методов ИНС. Бионические аспекты реализуемых сегодня ИНС-подходов.	1,2,4,9-15

4	20	Обзор современного программного обеспечения для анализа больших данных. Основные поставщики на рынках решений для управления большими данными.	8,9-15
---	----	--	--------

Методические указания по самостоятельному изучению отдельных разделов дисциплины приведены в соответствующем разделе ИОС.

10. Расчетно-графическая работа

Не предусмотрена

11. Курсовая работа

Не предусмотрена

12. Курсовой проект

Не планируется

13. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине (модулю)

Процедура оценивания знаний, умений и навыков проводится в соответствии со следующими методическими материалами и заключается в проведении устного экзаменационного опроса в виде диалога преподавателя со студентом, цель которого – систематизация и уточнение имеющихся у студента знаний, проверка его индивидуальных возможностей усвоения материала; отчетов по лабораторным работам, для оценки способности студента применить полученные ранее знания для организации системы управления информационной безопасностью, в проведении модулей и коллоквиумов, как способов межсессионной проверки знаний, умений, навыков студента в середине семестра по пройденным темам изучаемого предмета.

Показателем оценивания степени усвоения знаний является оценка, полученная на зачете при ответе на вопросы для экзамена. Оценка выставляется по четырехбалльной шкале, соответствующей оценкам «отлично», «хорошо», «удовлетворительно», «неудовлетворительно» и осуществляется путем анализа ответа на вопросы для экзамена. При этом руководствуются следующими критериями.

Оценка	Критерии оценивания результатов обучения (дескрипторы)
Отлично	заслуживает обучающийся, обнаруживший всестороннее, систематическое и глубокое знание учебного материала, предусмотренного программой, усвоивший основную литературу и знакомый с дополнительной литературой, рекомендованной программой.
Хорошо	заслуживает обучающийся, обнаруживший полное знание учебного материала, усвоивший основную литературу, рекомендованную в программе. Оценка "хорошо" выставляется обучающимся, показавшим систематический характер знаний по дисциплине и

	способным к их самостоятельному пополнению и обновлению в ходе дальнейшей учебной работы и профессиональной деятельности.
Удовлетворительно	заслуживает обучающийся, обнаруживший знания основного учебного материала в объеме, необходимом для дальнейшей учебы и предстоящей работы по профессии, знакомых с основной литературой, рекомендованной программой. Оценка выставляется обучающимся, допустившим погрешности в ответе на экзамене и при выполнении экзаменационных заданий, но обладающим необходимыми знаниями для их устранения под руководством преподавателя.
Неудовлетворительно	выставляется обучающемуся, обнаружившему пробелы в знаниях основного учебного материала. Оценка ставится обучающимся, которые не могут продолжить обучение или приступить к профессиональной деятельности по окончании образовательного учреждения без дополнительных занятий по рассматриваемой дисциплине.

Умения и навыки, приобретенные студентом на этапе освоения указанной части компетенций при преподавании рассматриваемой дисциплины, оцениваются по результатам выполнения практических заданий, включающих одну или несколько задач (вопросов) в виде краткой формулировки действий (комплекса действий), которые следует выполнить, или описание результата, который нужно получить. Показателем оценивания степени усвоения знаний является оценка, полученная при представлении материалов и докладе по выданной теме. Оценка выставляется по четырехбалльной шкале, соответствующей оценкам «отлично», «хорошо», «удовлетворительно», «неудовлетворительно» и осуществляется путем анализа представленного материала в ответ на практические контрольные задания. При этом руководствуются следующими критериями:

Оценка	Критерии оценивания результатов обучения (дескрипторы)
Отлично	5 баллов выставляется студенту, если задание выполнено в полном объеме с соблюдением необходимой последовательности. Студенты работают полностью самостоятельно: подбирают необходимые для выполнения предлагаемых работ в задании источники знаний, показывают необходимые для проведения практической работы теоретические знания, практические умения и навыки.
Хорошо	4 балла выставляется студенту, если задание выполнено в полном объеме и самостоятельно. Допускаются отклонения от необходимой последовательности

	выполнения, не влияющие на правильность конечного результата. Студенты используют указанные преподавателем источники знаний, включая страницы атласа, таблицы из приложения к учебнику, страницы из справочной литературы по предмету. Задание показывает знание учащихся основного теоретического материала и овладение умениями, необходимыми для самостоятельного выполнения работы. Могут быть неточности и небрежность в оформлении результатов работы.
удовлетворительно	3 балла выставляется студенту, если задание на практическую работу выполняется и оформляется студентами при помощи преподавателя или хорошо подготовленных и уже выполненных на «отлично» данную работу студентов. На выполнение задания затрачивается много времени (можно дать возможность доделать работу дома). Студенты показывают знания теоретического материала, но испытывают затруднение при самостоятельной работе с физическими приборами, графиками, таблицами справочной литературы.
неудовлетворительно	2 балла выставляется, если студенты показывают плохое знание теоретического материала и отсутствие умения применить знания к решению практической задачи. Руководство и помощь со стороны преподавателя и хорошо подготовленных студентов неэффективны по причине плохой подготовки студента.

Процедура оценивания знаний, умений, навыков включает учет успешности выполнения практических работ, самостоятельной работы и сдачу экзамена.

Критерии сформированности компетенций:

При проверке знаний по компетенции в целом, положительное решение о сформированности компетенции принимается в случае правильного ответа не менее чем 30 % вопросов теста и/или ответа на поставленные на зачете вопросы в соответствии с указанными выше уровнями освоения компетенций, при условии выполнения полного комплекса практических работ по дисциплине, отчета по каждой и выполнения заданий на самостоятельную работу студента.

При проверке умения в соответствии с уровнями освоения компетенции студенту предоставляется возможность после выполнения практической работы и заданий на самостоятельную работу, предоставления отчета по практической работе в соответствии с требованиями, представленными в методических указаниях по практической работе, отчета по практической работе, продемонстрировать действия по использованию методов и средств в объеме уровней освоения.

При проверке владения навыковыми составляющими компетенций студенту предоставляется возможность решения профессиональных задач в соответствии с уровнями их освоения, с оценкой полноты предлагаемых для решения методов, оптимальности выбора метода и средств ее решения, устойчивости демонстрируемых способностей по выполнению действий в соответствии с уровнями освоения компетенции.

Вопросы для зачета

Не предусмотрены

Вопросы для экзамена

1. Истоки возникновения проблемы обработки больших данных
2. Области использования больших данных. Особенности, характерные для больших данных, и их дифференциация от классических данных.
3. Принципы адаптации алгоритмов анализа больших данных.
4. Оценка перспектив развития теории и прикладных методик для анализа больших данных в различных областях естественно-научного знания.
5. Методики анализа больших данных: A/B-тестирование.
6. Методики анализа больших данных: Регрессионный анализ.
7. Методики анализа больших данных: Association rule learning.
8. Методики анализа больших данных: Cluster analysis.
9. Методики анализа больших данных: Classification.
10. Методики анализа больших данных: Crowdsourcing.
11. Методики анализа больших данных: Data mining.
12. Проблема адаптивных алгоритмов анализа в плохо структурированных данных и пути её решения.
13. Принципы построения ИНС. Классификация ИНС.
14. Принципы построения ИНС. Различные методы обучения ИНС.
15. Подходы к анализу больших данных на базе генетических алгоритмов (Genetic algorithms) и самообучения ИНС (Machine learning).
16. Принципы построения информационных систем и систем распознавания образов для анализа больших данных.
17. Структура современной информационной системы автоматизированного технологического оборудования.
18. Проблема визуализации результатов анализа больших данных и их дальнейшая постобработка.

14. Образовательные технологии

Для достижения планируемых результатов используются следующие образовательные технологии:

1. Дистанционные на основе информационно-образовательной среды СГТУ имени Гагарина Ю.А., основе реализации возможности самостоятельного изучения материалов по всем видам образовательной деятельности в соответствии с учебным планом, в том числе до прохождения занятий, текущего дистанционного консультирования студентов.

2. Развивающее проблемно-ориентированное обучение, направленное на формирование и развитие проблемного мышления, мыслительной активности, способности видеть и формулировать проблемы, выбирать способы и средства для их решения на основе рассмотрения примеров из практической деятельности преподавателей, в области научно-практических исследований.

3. Личностно ориентированное обучение, обеспечивающие в ходе учебного процесса учет различных способностей обучаемых, создание необходимых условий для развития их индивидуальных способностей, развитие активности личности в учебном процессе в рамках самостоятельной работы.

15. ПЕРЕЧЕНЬ УЧЕБНО-МЕТОДИЧЕСКОГО ОБЕСПЕЧЕНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ

Обязательные издания

1. Воронова Л.И. Big Data. Методы и средства анализа [Электронный ресурс] : учебное пособие / Л.И. Воронова, В.И. Воронов. — Электрон. текстовые данные. — М. : Московский технический университет связи и информатики, 2016. — 33 с. — 2227-8397. — Режим доступа: <http://www.iprbookshop.ru/61463.html>
2. Горожанина Е.И. Нейронные сети [Электронный ресурс] : учебное пособие / Е.И. Горожанина. — Электрон. текстовые данные. — Самара: Поволжский государственный университет телекоммуникаций и информатики, 2017. — 84 с. — 2227-8397. — Режим доступа: <http://www.iprbookshop.ru/75391.html>
3. Медведев Д.М. Структуры и алгоритмы обработки данных в системах автоматизации и управления [Электронный ресурс] : учебное пособие / Д.М. Медведев. — Электрон. текстовые данные. — Саратов: Ай Пи Эр Медиа, 2018. — 100 с. — 978-5-4486-0192-7. — Режим доступа: <http://www.iprbookshop.ru/71591.html>
4. Самуйлов С.В. Алгоритмы и структуры обработки данных [Электронный ресурс] : учебное пособие / С.В. Самуйлов. — Электрон. текстовые данные. — Саратов: Вузовское образование, 2016. — 132 с. — 2227-8397. — Режим доступа: <http://www.iprbookshop.ru/47275.html>
5. Белаш В.Ю. Моделирование потоков данных в информационных системах [Электронный ресурс] : учебное пособие / В.Ю. Белаш, Н.В. Тимошина. — Электрон. текстовые данные. — Саратов: Вузов-

ское образование, 2018. — 58 с. — 978-5-4487-0256-3. — Режим доступа: <http://www.iprbookshop.ru/75683.html>

Дополнительные издания

6. Яхьяева Г.Э. Нечеткие множества и нейронные сети [Электронный ресурс] : учебное пособие / Г.Э. Яхьяева. — Электрон. текстовые данные. — Москва, Саратов: Интернет-Университет Информационных Технологий (ИНТУИТ), Вузовское образование, 2017. — 320 с. — 978-5-4487-0079-8. — Режим доступа: <http://www.iprbookshop.ru/67390.html>
7. Селиванова И.А. Построение и анализ алгоритмов обработки данных [Электронный ресурс] : учебно-методическое пособие / И.А. Селиванова, В.А. Блинов. — Электрон. текстовые данные. — Екатеринбург: Уральский федеральный университет, ЭБС АСВ, 2015. — 108 с. — 978-5-7996-1489-8. — Режим доступа: <http://www.iprbookshop.ru/68277.html>
8. Курапова Е.В. Структуры и алгоритмы обработки данных [Электронный ресурс] : лабораторный практикум / Е.В. Курапова, Е.П. Мачикина. — Электрон. текстовые данные. — Новосибирск: Сибирский государственный университет телекоммуникаций и информатики, 2015. — 23 с. — 2227-8397. — Режим доступа: <http://www.iprbookshop.ru/55501.html>
9. Введение в математический пакет Matlab [Электронный ресурс] : учебно-методическое пособие / . — Электрон. текстовые данные. — М. : Московский технический университет связи и информатики, 2016. — 88 с. — 2227-8397. — Режим доступа: <http://www.iprbookshop.ru/61469.html>

Периодические издания

10. Современные технологии автоматизации -
Режим доступа: http://elibrary.ru/title_about.asp?id=9119
11. Вестник СГТУ -
Режим доступа: <http://lib.sstu.ru/index.php/elmrazdel/melellib/91-mperiodizdan>
12. Автоматизация. Современные технологии -
Режим доступа: http://elibrary.ru/title_about.asp?id=7647
13. Мехатроника, автоматизация, управление -
Режим доступа: http://elibrary.ru/title_about.asp?id=8851

Интернет ресурсы

14. <http://www.aktakom.ru> – Современные измерительные приборы и датчики
15. <http://www.ncsystems.ru> – Компьютерные системы управления

Источники ИОС:

16.

<https://portal3.sstu.ru/Facult/INETM/AUM/15.03.04/%D0%91.1.3.4.1/default.aspx>

16. Материально-техническое обеспечение

Перечень и описание учебных аудиторий (специализированная учебная мебель, мультимедиа и наборы учебно-наглядных пособий, соответствующие примерным программам дисциплин и УМК):

- типовая лекционная аудитория, оборудована соответствующей учебной мебелью, мультимедийным проектором, экраном, доской для записей фломастером или мелом и наглядными пособиями в виде плакатов;
- типовая лабораторная аудитория оборудована доской для записей фломастером или мелом, стендами для лабораторных работ по дисциплине ТСА, 2 компьютера для обработки результатов экспериментов.

Перечень и описание помещений для самостоятельной работы: компьютерный класс каф. АУМ (компьютеры с выходом в Интернет) –

Информационное и учебно-методическое обеспечение:

Сайт СГТУ ИОС по ссылке:

<https://portal3.sstu.ru/Facult/INETM/AUM/15.03.04/B.2.3.1.2/default.aspx>

(раздел «Учебные материалы» в папке «2.2 Методические указания по выполнению лабораторных работ»).

Электронно-библиотечная система, электронная библиотека вуза:

<http://lib.sstu.ru/index.php/elmrzdel/melellib>;

Перечень оборудования информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень программного обеспечения и информационных справочных систем (при необходимости):

- персональный компьютер;
- мультимедийный проектор;
- ПО Microsoft Power Point 2007;

Программно-технические средства, используемые при выполнении лабораторных работ:

- персональные компьютеры в составе ЛВС кафедры АУМ;
- Microsoft Office 2007;
- ПО для статистического анализа данных в среде Matlab.