Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Максимов Алексей Борисович
Должность: директор департамента по образовательной политике
Дата подписания: 23.10.2023 16:02:46
Уникальный программный ключ:
8db180d1a3f02ac9e60521a5672742735c18b1d6

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное автономное образовательное учреждение высшего образования
«МОСКОВСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

УТВЕРЖДАЮ

Декан факультета информационных технологий

/Д. Г. Демидов/

«28» 04 2022 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

«Хранилища данных»

Направление подготовки
09.03.03 Прикладная информатика
Образовательная программа (профиль подготовки)
«Большие и открытые данные»

Квалификация (степень) выпускника
бакалавр

Форма обучения
очная

Москва 2022

Программа дисциплины «Хранилища данных» составлена в соответствии с требованиями ФГОС ВО и учебным планом по направлению **09.03.03 «Прикладная информатика»** и профилю подготовки **«Большие и открытые данные».**

Программу составил 

_____/А.В. Осипов/

Программа дисциплины утверждена на заседании кафедры «Прикладная информатика»
«____» августа 2022 г. протокол № _____

Заведующий кафедрой
доцент, к.э.н.  /С.В.Суворов/

Программа согласована с руководителем образовательной программы по направлению подготовки **09.03.03 «Прикладная информатика»** по профилю подготовки **«Большие и открытые данные».**


_____ /С.В.Суворов/
«____» августа 2022 г.

Программа утверждена на заседании учебно-методической комиссии факультета Информационных технологий

Председатель комиссии  / Д. Г. Демидов/

«____»_____2022 г. Протокол:

1. Цели освоения дисциплины.

К **основным целям** освоения дисциплины «Хранилище данных» следует отнести:

– расширенное формирование у студентов представления о принципах и методах машинного обучения;

– знакомство студентов с современными методами работы с большими данными.

К **основным задачам** освоения дисциплины «Хранилище данных» следует отнести:

– освоение методологии обработки больших данных;

– использование компьютерных технологий реализации методов машинного обучения.

2. Место дисциплины в структуре ООП магистратуры

Дисциплина «Хранилище данных» относится к вариативной части основной образовательной программы бакалавриата.

Она взаимосвязана логически и содержательно-методически со следующими дисциплинами ООП:

- База данных;
- Программирование;
- Интеллектуальный анализ данных.

Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенные с планируемыми результатами освоения образовательной программы

В результате освоения дисциплины (модуля) у обучающихся формируются следующие компетенции и должны быть достигнуты следующие результаты обучения как этап формирования соответствующих компетенций:

Код компетенции	**В результате освоения образовательной программы обучающийся должен обладать**	**Перечень планируемых результатов обучения по дисциплине**
ОПК-3	способностью к разработке алгоритмических и программных решений в области системного и прикладного программирования, математических, информационных и имитационных моделей, созданию информационных ресурсов глобальных сетей, образовательного контента, прикладных баз данных, тестов и средств тестирования систем и средств на соответствие стандартам и исходным требованиям	**знать:** • основные методы проектирования и эксплуатации хранилищ данных и технологий **уметь:** • использовать хранилища данных и технологии для решения научных задач **владеть:** • методами проектирования и эксплуатации хранилищ данных и технологий для решения практических научных задач
ПК-2	способностью понимать, совершенствовать и применять современный математический аппарат	**знать:** • концептуальные и теоретические подходы к построению хранилищ данных и использованию технологий **уметь:** • использовать методы проектирования и эксплуатации хранилищ данных и технологий для решения научных проблем и задач **владеть:** • методами проектирования и эксплуатации хранилищ данных и

		технологий для проведения научных исследований

4. Структура и содержание дисциплины

Общая трудоемкость дисциплины составляет 9 зачетных единиц, т.е. 324 академических часов (из них 180 часов – самостоятельная работа студентов): лекции – 36 часов, лабораторные работы – 108 часа, форма контроля – зачет

Разделы дисциплины «Хранилище данных» изучаются на четвертом курсе.

Структура и содержание дисциплины «Хранилище данных» по видам работы отражены в приложении.

Содержание разделов дисциплины

Седьмой семестр

1. Область применения хранилищ данных и больших данных и их преимущества
2. Виды данных
2.1 Структурированные данные
2.2 Неструктурированные данные
2.3 Данные на естественном языке
2.4 Машинные данные
2.5 Графовые, или сетевые, данные
2.6 Аудио, видео и графика
2.7 Потоковые данные
3. Экосистема больших данных
3.1 Распределенные файловые системы
3.2 Инфраструктура распределенного программирования
3.3 Инфраструктура интеграции данных
3.4 Инфраструктуры машинного обучения
3.5 Базы данных NoSQL

Восьмой семестр

4. Технологии Data Mining
4.1 Задачи Data Mining

Обзор задач Data Mining. Стандартизация подхода к решению задач Data Mining. Процесс CRISP-DM. Виды данных. Кластеризация, классификация, регрессия. Понятие модели и алгоритма обучения.

4.2 Задача кластеризации и EM-алгоритм

Постановка задачи кластеризации. Функции расстояния. Критерии качества кластеризации. EM-алгоритм. K-means и модификации.

4.3 Различные алгоритмы кластеризации

Иерархическая кластеризация. Agglomerative и Divisive алгоритмы. Различные виды расстояний между кластерами. Stepwise-optimal алгоритм. Случай неэвклидовых пространств. Критерии выбора количества кластеров: rand, silhouette. DBSCAN.

4.4 Задача классификации
Постановка задач классификации и регрессии. Теория принятия решений. Виды моделей. Примеры функций потерь. Переобучение. Метрики качества классификации. MDL. Решающие деревья. Алгоритм CART.
4.5 Обработка текстов, Naive Bayes
Условная вероятность и теорема Байеса. Нормальное распределение. Naive Bayes: multinomial, binomial, gaussian. Сглаживание. Генеративная модель NB и байесовский вывод. Графические модели.
4.6 Линейные модели для классификации и регрессии
Обобщенные линейные модели. Постановка задачи оптимизации. Примеры критериев. Градиентный спуск. Регуляризация. Метод Maximum Likelihood. Логистическая регрессия.
4.7 Машина опорных векторов
Разделяющая поверхность с максимальным зазором. Формулировка задачи оптимизации для случаев линейно-разделимых и линейно-неразделимых классов. Сопряженная задача. Опорные векторы. KKT-условия. SVM для задач классификации и регрессии. Kernel trick. Теорема Мерсера. Примеры функций ядра.
4.8 Методы снижения размерности пространства
Проблема проклятия размерности. Отбор и выделение признаков. Методы выделения признаков (feature extraction). Метод главных компонент (PCA). Метод независимых компонент (ICA). Методы основанные на автоэнкодерах. Методы отбора признаков (feature selection). Методы основанные на взаимной корреляции признаков. Метод максимальной релевантность и минимальной избыточности (mRMR). Методы основанные на деревьях решений.
4.9 Алгоритмические композиции
Комбинации классификаторов. Модельные деревья решений. Смесь экспертов. Stacking. Стохастические методы построения ансамблей классификаторов. Bagging. RSM. Алгоритм RandomForest.
4.10 Глубокие нейронные сети
Трудности обучения многослойного персептрона. Предобучение используя РБМ. Глубокий автоэнкодер, глубокая многослойная нейросеть. Deep belief network и deep Boltzmann machine. Сверточные сети.

5. Образовательные технологии.

Методика преподавания дисциплины «Хранилище данных» и реализация компетентного подхода в изложении и восприятии материала предусматривает использование следующих активных и интерактивных форм проведения групповых, индивидуальных, аудиторных занятий в сочетании с внеаудиторной работой с целью формирования и развития профессиональных навыков обучающихся:

– подготовка к выполнению практических работ в компьютерных классах вуза;

– подготовка, представление и обсуждение презентаций на семинарских занятиях;

– организация и проведение интерактивных форм текущего контроля знаний студентов в форме выполнения индивидуальных заданий.

Удельный вес занятий, проводимых в интерактивных формах, определен главной целью образовательной программы, особенностью контингента обучающихся и содержанием

дисциплины «Хранилище данных» и в целом по дисциплине составляет 50% аудиторных занятий. Занятия лекционного типа составляют 50% от объема аудиторных занятий.

6. Оценочные средства для текущего контроля успеваемости, промежуточной аттестации по итогам освоения дисциплины и учебно-методическое обеспечение самостоятельной работы студентов

В процессе обучения используются следующие оценочные формы самостоятельной работы студентов, оценочные средства текущего контроля успеваемости и промежуточных аттестаций:

- индивидуальное задание на работу с хранилищами данных и большими данными;
- подготовка к выполнению практических работ и их защита.

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

6.1.1. Перечень компетенций с указанием этапов их формирования в процессе освоения образовательной программы.

В результате освоения дисциплины (модуля) формируются следующие компетенции:

Код компетенции	В результате освоения образовательной программы обучающийся должен обладать
ОПК-18	способностью к разработке алгоритмических и программных решений в области системного и прикладного программирования, математических, информационных и имитационных моделей, созданию информационных ресурсов глобальных сетей, образовательного контента, прикладных баз данных, тестов и средств тестирования систем и средств на соответствие стандартам и исходным требованиям
ПК-2	способностью понимать, совершенствовать и применять современный математический аппарат

В процессе освоения образовательной программы данные компетенции, в том числе их отдельные компоненты, формируются поэтапно в ходе освоения обучающимися дисциплин (модулей), практик в соответствии с учебным планом и календарным графиком учебного процесса.

6.1.2. Описание показателей и критериев оценивания компетенций, формируемых по итогам освоения дисциплины (модуля), описание шкал оценивания

Показателем оценивания компетенций на различных этапах их формирования является достижение обучающимися планируемых результатов обучения по дисциплине (модулю).

ОПК-3 способностью к разработке алгоритмических и программных решений в области системного и прикладного программирования, математических, информационных и имитационных моделей, созданию информационных ресурсов глобальных сетей, образовательного контента, прикладных баз данных, тестов и средств тестирования систем и средств на соответствие стандартам и исходным требованиям				
Показатель	**Критерии оценивания**			
	2	**3**	**4**	**5**
знать: основные методы проектирования и эксплуатации хранилищ данных и технологий	Обучающийся демонстрирует полное отсутствие или недостаточное соответствие следующих знаний: основные методы проектирования и эксплуатации хранилищ данных и технологий	Обучающийся демонстрирует неполное соответствие следующих знаний: основные методы проектирования и эксплуатации хранилищ данных и технологий. Допускаются значительные ошибки, проявляется недостаточность знаний, по ряду методов, обучающийся испытывает значительные затруднения при оперировании знаниями при их переносе на новые ситуации.	Обучающийся демонстрирует частичное соответствие следующих знаний: основные методы проектирования и эксплуатации хранилищ данных и технологий, но допускаются незначительные ошибки, неточности, затруднения при рассмотрении новых ситуаций.	Обучающийся демонстрирует полное соответствие следующих знаний: основные методы проектирования и эксплуатации хранилищ данных и технологий, свободно оперирует приобретенными знаниями.

уметь: использовать методы проектирования и эксплуатации хранилищ данных и технологий для решения научных задач	Обучающийся не умеет или в недостаточной степени умеет использовать методы проектирования и эксплуатации хранилищ данных и технологий для решения научных задач	Обучающийся демонстрирует неполное соответствие следующих умений: использовать методы проектирования и эксплуатации хранилищ данных и технологий для решения научных задач. Допускаются значительные ошибки, проявляется недостаточность умений, обучающийся испытывает значительные затруднения при оперировании умениями при их переносе на новые ситуации.	Обучающийся демонстрирует частичное соответствие следующих умений: использовать методы проектирования и эксплуатации хранилищ данных и технологий для решения научных задач. Умения освоены, но допускаются незначительные ошибки, неточности, затруднения при переносе умений на новые, нестандартные ситуации.	Обучающийся демонстрирует полное соответствие следующих умений: использовать методы проектирования и эксплуатации хранилищ данных и технологий для решения научных задач. Свободно оперирует приобретенными умениями, применяет их в ситуациях повышенной сложности.
владеть: методами проектирования и эксплуатации хранилищ данных и технологий для решения практических научных задач	Обучающийся не владеет или в недостаточной степени владеет методами проектирования и эксплуатации хранилищ данных и технологий для решения практических научных задач	Обучающийся владеет методами проектирования и эксплуатации хранилищ данных и технологий для решения практических научных задач в неполном объеме,	Обучающийся частично владеет методами проектирования и эксплуатации хранилищ данных и технологий для решения практических научных задач,	Обучающийся в полном объеме владеет методами проектирования и эксплуатации хранилищ данных и технологий для решения практических научных задач, свободно применяет полученные навыки в ситуациях повышенной сложности.

		допускаются значительные ошибки, проявляется недостаточность владения навыками по ряду задач. Обучающийся испытывает значительные затруднения при применении навыков в новых ситуациях.	навыки освоены, но допускаются незначительные ошибки, неточности, затруднения при переносе умений на новые, нестандартные ситуации.	
ПК-2 способностью понимать, совершенствовать и применять современный математический аппарат				
знать: концептуальные и теоретические подходы к проектированию и эксплуатации хранилищ данных и технологий	Обучающийся демонстрирует полное отсутствие или недостаточное соответствие следующих знаний: концептуальные и теоретические подходы к проектированию и эксплуатации хранилищ данных и технологий	Обучающийся демонстрирует неполное соответствие следующих знаний: концептуальные и теоретические подходы к проектированию и эксплуатации хранилищ данных и технологий. Допускаются значительные ошибки, проявляется недостаточность знаний, обучающийся испытывает значительные затруднения при	Обучающийся демонстрирует частичное соответствие следующих знаний: концептуальные и теоретические подходы к проектированию и эксплуатации хранилищ данных и технологий, но допускаются незначительные ошибки, неточности, затруднения при рассмотрении новых ситуаций.	Обучающийся демонстрирует полное соответствие следующих знаний: концептуальные и теоретические подходы к проектированию и эксплуатации хранилищ данных и технологий, свободно оперирует приобретенными знаниями.

		оперировании знаниями при их переносе на новые ситуации.		

уметь: использовать методы проектирования и эксплуатации хранилищ данных и технологий Big Data для решения научных проблем и задач	Обучающийся не умеет или в недостаточной степени умеет использовать методы проектирования и эксплуатации хранилищ данных и технологий Big Data для решения научных проблем и задач	Обучающийся демонстрирует неполное соответствие следующих умений: использовать методы проектирования и эксплуатации хранилищ данных и технологий Big Data для решения научных проблем и задач. Допускаются значительные ошибки, проявляется недостаточность умений, обучающийся испытывает значительные затруднения при оперировании умениями при их переносе на новые ситуации.	Обучающийся демонстрирует частичное соответствие следующих умений: использовать методы проектирования и эксплуатации хранилищ данных и технологий Big Data для решения научных проблем и задач Умения освоены, но допускаются незначительные ошибки, неточности, затруднения при переносе умений на новые, нестандартные ситуации.	Обучающийся демонстрирует полное соответствие следующих умений: использовать методы проектирования и эксплуатации хранилищ данных и технологий Big Data для решения научных проблем и задач. Свободно оперирует приобретенными умениями, применяет их в ситуациях повышенной сложности.
владеть: методами проектирования и эксплуатации хранилищ данных и технологий Big Data для проведения научных	Обучающийся не владеет или в недостаточной степени владеет методами проектирования и эксплуатации хранилищ данных и технологий Big Data для проведения	Обучающийся владеет методами проектирования и эксплуатации хранилищ данных и технологий Big Data для проведения научных исследований в	Обучающийся частично владеет методами проектирования и эксплуатации хранилищ данных и технологий Big Data для проведения	Обучающийся в полном объеме владеет методами проектирования и эксплуатации хранилищ данных и технологий Big Data для проведения научных исследований, свободно применяет полученные навыки в ситуациях повышенной сложности.

исследований	научных исследований	неполном объеме, допускаются значительные ошибки. Обучающийся испытывает значительные затруднения при применении навыков в новых ситуациях.	научных исследований, навыки освоены, но допускаются незначительные ошибки, неточности, затруднения при переносе умений на новые, нестандартные ситуации.	

Шкалы оценивания результатов промежуточной аттестации и их описание:

Форма промежуточной аттестации: зачет.

Промежуточная аттестация обучающихся в форме зачёта проводится по результатам выполнения всех видов учебной работы, предусмотренных учебным планом по данной дисциплине (модулю), при этом учитываются результаты текущего контроля успеваемости в течение семестра. Оценка степени достижения обучающимися планируемых результатов обучения по дисциплине (модулю) проводится преподавателем, ведущим занятия по дисциплине (модулю) методом экспертной оценки. По итогам промежуточной аттестации по дисциплине (модулю) выставляется оценка «зачтено» или «не зачтено».

Шкала оценивания	**Описание**
Зачтено	Выполнены все виды учебной работы, предусмотренные учебным планом. Студент демонстрирует соответствие знаний, умений, навыков приведенным в таблицах показателей, оперирует приобретенными знаниями, умениями, навыками, применяет их в ситуациях повышенной сложности. При этом могут быть допущены незначительные ошибки, неточности, затруднения при аналитических операциях, переносе знаний и умений на новые, нестандартные ситуации.
Не зачтено	Не выполнен один или более видов учебной работы, предусмотренных учебным планом. Студент демонстрирует неполное соответствие знаний, умений, навыков приведенным в таблицах показателей, допускаются значительные ошибки, проявляется отсутствие знаний, умений, навыков по ряду показателей, студент испытывает значительные затруднения при оперировании знаниями и умениями при их переносе на новые ситуации.

Форма промежуточной аттестации: экзамен.

Промежуточная аттестация обучающихся в форме экзамена проводится по результатам выполнения всех видов учебной работы, предусмотренных учебным планом по данной дисциплине (модулю), при этом учитываются результаты текущего контроля успеваемости в течение семестра. Оценка степени достижения обучающимися планируемых результатов обучения по дисциплине (модулю) проводится преподавателем, ведущим занятия по дисциплине (модулю) методом экспертной оценки. По итогам промежуточной аттестации по дисциплине (модулю) выставляется оценка «отлично», «хорошо», «удовлетворительно» или «неудовлетворительно».

Шкала оценивания	***Описание***
Отлично	*Выполнены все виды учебной работы, предусмотренные учебным планом. Студент демонстрирует соответствие знаний, умений, навыков приведенным в таблицах показателей,*

	оперирует приобретенными знаниями, умениями, навыками, применяет их в ситуациях повышенной сложности. При этом могут быть допущены незначительные ошибки, неточности, затруднения при аналитических операциях, переносе знаний и умений на новые, нестандартные ситуации.
Хорошо	*Выполнены все виды учебной работы, предусмотренные учебным планом. Студент демонстрирует неполное, правильное соответствие знаний, умений, навыков приведенным в таблицах показателей, либо если при этом были допущены 2-3 несущественные ошибки.*
Удовлетворительно	*Выполнены все виды учебной работы, предусмотренные учебным планом. Студент демонстрирует соответствие знаний, в котором освещена основная, наиболее важная часть материала, но при этом допущена одна значительная ошибка или неточность.*
Неудовлетворительно	*Не выполнен один или более видов учебной работы, предусмотренных учебным планом. Студент демонстрирует неполное соответствие знаний, умений, навыков приведенным в таблицах показателей, допускаются значительные ошибки, проявляется отсутствие знаний, умений, навыков по ряду показателей, студент испытывает значительные затруднения при оперировании знаниями и умениями при их переносе на новые ситуации.*

Фонды оценочных средств представлены в приложении 1 к рабочей программе.

Приложение 1 к
рабочей программе

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное автономное образовательное учреждение
высшего образования
«МОСКОВСКИЙ ПОЛИТЕХНИЧЕСКИЙ
УНИВЕРСИТЕТ»

Направление подготовки
09.03.03 «Прикладная информатика»

Квалификация (степень) выпускника
Бакалавр

Кафедра: Прикладная информатика

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

Хранилище данных

Состав: 1. Паспорт фонда оценочных средств

2. Описание оценочных средств:
Кейсовые задания на использование методов хранилищ данных

3. Контрольные вопросы

4. Экзаменационные вопросы

Москва, 2022 год

1. Паспорт фонда оценочных средств

Оценочные средства включают:
- кейсовые задания на использование методов хранилищ данных и Big Data
.

2. Описание оценочных средств:

Пример кейсового задания: Прогнозирование вредоносных URL-адресов
Загрузите данные по адресу http://sysnet.ucsd.edu/projects/url/#datasets и поместите их в рабочий каталог. Выберите данные в формате SVMLight.
Задача: создать стохастический градиентный классификатор сайтов. Целевая переменная может быть равна 1 или -1. "1": сайт безопасен для посещения, "-1": сайт небезопасен.
Результы выполнения оформить в виде:
Этап 1: Определение цели исследования
Этап 2: Сбор данных URL
Этап 3: Исследование данных
Этап 4: Построение модели

Пример кейсового задания: Построение рекомендательной системы внутри базы данных
Рекомендательная система должна выдавать результаты вида: «Клиенты, смотревшие этот фильм, также смотрели…»
Для обучения использовать метод “k-ближайших соседей“.
Исходные данные для построения рекомендательной системы создаются самими учащимися.
Данные должны быть вида:

Клиент	Фильм 1	Фильм 2	Фильм 3	...	Фильм 32
Jack Dani	1	0	0		1
Wilhelmson	1	1	0		1
...					
Jane Dane	0	0	1		0
Xi Liu	0	0	0		1
Eros Mazo	1	1	0		1
...					

Результаты выполнения оформить в виде:
Этап 1: Вопрос исследования
Этап 2: Подготовка данных
Этап 3: Построение модели
Этап 4: Отображение и автоматизация

3. Контрольные вопросы

1. Область применения хранилищ данных и больших данных и их преимущества
2. Виды данных. Структурированные данные
3. Виды данных. Неструктурированные данные
4. Виды данных. Данные на естественном языке
5. Виды данных. Машинные данные

6. Виды данных. Графовые, или сетевые, данные
7. Виды данных. Аудио, видео и графика
8. Виды данных. Потоковые данные
9. Понятие гиперкуба. OLAP.
10. ROLAP-системы
11. MOLAP-системы
12. HOLAP-системы
13. Распределенные файловые системы
14. Базы данных NoSQL

4. Экзаменационные вопросы

1. Обзор задач Data Mining. Кластеризация, классификация, регрессия.
2. Постановка задачи кластеризации. Функции расстояния. Критерии качества кластеризации.
3. EM-алгоритм.
4. K-means и модификации.
5. Иерархическая кластеризация. Agglomerative и Divisive алгоритмы.
6. Постановка задач классификации и регрессии. Теория принятия решений.
7. Виды моделей. Примеры функций потерь. Переобучение. Метрики качества классификации.
8. MDL. Решающие деревья.
9. Алгоритм CART.
10. Условная вероятность и теорема Байеса. Нормальное распределение.
11. Naive Bayes: multinomial, binomial, gaussian. Сглаживание. Генеративная модель NB и байесовский вывод.
12. Обобщенные линейные модели. Постановка задачи оптимизации. Примеры критериев. Градиентный спуск.
13. Разделяющая поверхность с максимальным зазором. Формулировка задачи оптимизации для случаев линейно-разделимых и линейно-неразделимых классов.
14. KKT-условия. SVM для задач классификации и регрессии.
15. Проблема проклятия размерности. Отбор и выделение признаков. Методы выделения признаков (feature extraction).
16. Метод главных компонент (PCA).
17. Метод независимых компонент (ICA).
18. Методы, основанные на автоэнкодерах.
19. Комбинации классификаторов. Модельные деревья решений.
20. Алгоритм RandomForest.
21. Глубокий автоэнкодер, глубокая многослойная нейросеть.
22. Deep belief network и deep Boltzmann machine.
23. Сверточные сети.

ПОКАЗАТЕЛЬ УРОВНЯ СФОРМИРОВАННОСТИ КОМПЕТЕНЦИЙ

Хранилища данных
ФГОС ВО 09.03.03 «Прикладная информатика»
В процессе освоения данной дисциплины студент формирует и демонстрирует следующие общепрофессиональные и **профессиональные компетенции:**

КОМПЕТЕНЦИИ		**Перечень компонентов**	**Технология формирования компетенций**	**Форма оценочного средства****	**Степени уровней освоения компетенций**
ИНДЕКС	**ФОРМУЛИРОВКА**				
ОПК-3	способностью к разработке алгоритмических и программных решений в области системного и прикладного программирования, математических, информационных и имитационных моделей, созданию информационных ресурсов глобальных сетей, образовательного контента, прикладных баз данных, тестов и средств тестирования систем и средств на соответствие стандартам и исходным требованиям	**знать:** • основные методы проектирования и эксплуатации хранилищ данных и технологий **уметь:** • использовать методы проектирования и эксплуатации хранилищ данных и технологий для решения научных задач **владеть:** • методами проектирования и эксплуатации хранилищ данных и технологий для решения практических научных задач	лекция, самостоятельная работа, семинарские занятия	ДИ, К-З, К, К/Р, РТ, РГР	**Базовый уровень** - способен использовать основные методы естественнонаучных дисциплин в профессиональной деятельности для экспериментального исследования **Повышенный уровень** **-** способен использовать основные методы естественнонаучных дисциплин в профессиональной деятельности для теоретического и экспериментального исследования

КОМПЕТЕНЦИИ		Перечень компонентов	Технология формирования компетенций	Форма оценочного средства**	Степени уровней освоения компетенций
ИНДЕКС	ФОРМУЛИРОВКА				
ПК-2	способностью понимать, совершенствовать и применять современный математический аппарат	**знать:** • концептуальные и теоретические подходы к проектированию и эксплуатации хранилищ данных и технологий **уметь:** • использовать методы проектирования и эксплуатации хранилищ данных и технологий для решения научных проблем и задач **владеть:** • методами проектирования и эксплуатации хранилищ данных и технологий для проведения научных исследований	лекция, самостоятельная работа, семинарские занятия	ДИ, К-З, К, К/Р, РТ, РГР	**Базовый уровень:** использовать соответствующий математический аппарат и инструментальные средства для обработки информации по теме исследования **Повышенный уровень:** использовать соответствующий математический аппарат и инструментальные средства для обработки, анализа и систематизации информации по теме исследования

**- Сокращения форм оценочных средств см. в приложении 2 к РП

.

Приложение 2
к рабочей программе

Перечень оценочных средств по дисциплине Хранилище данных

№ ОС	Наименование оценочного средства	Краткая характеристика оценочного средства	Представление оценочного средства в ФОС
1	Деловая и/или ролевая игра (ДИ)	Совместная деятельность группы обучающихся и педагогического работника под управлением педагогического работника с целью решения учебных и профессионально - ориентированных задач путем игрового моделирования реальной проблемной ситуации. Позволяет оценивать умение анализировать и решать типичные профессиональные задачи.	Тема (проблема), концепция, роли и ожидаемый результат по каждой игре
2	Кейс-задача (К-З)	Проблемное задание, в котором обучающемуся предлагают осмыслить реальную профессионально-ориентированную ситуацию, необходимую для решения данной проблемы.	Задания для решения кейс-задачи
3	Коллоквиум (К)	Средство контроля усвоения учебного материала темы, раздела или разделов дисциплины, организованное как учебное занятие в виде собеседования педагогического работника с обучающимися.	Вопросы по темам/разделам дисциплины
4	Контрольная работа (К/Р)	Средство проверки умений применять полученные знания для решения задач определенного типа по теме или разделу	Комплект контрольных заданий по вариантам
5	Рабочая тетрадь (РТ)	Дидактический комплекс, предназначенный для самостоятельной работы обучающегося и позволяющий оценивать уровень усвоения им учебного материала.	Образец рабочей тетради
6	Расчетно-графическая работа (РГР)	Средство проверки умений применять полученные знания по заранее определенной методике для решения задач или заданий по модулю или дисциплине в целом.	Комплект заданий для выполнения расчетно-графической работы

7. Учебно-методическое и информационное обеспечение дисциплины.

а) основная литература:

Паклин Н.Б., Орешков В.И. Бизнес-аналитика: от данных к знаниям [Электронный ресурс], СПб, Питер, 2013. – 704 стр. Режим доступа: http://www.padaread.com/?book=52105

б) дополнительная литература

Заенцев И.В. Нейронные сети: основные модели [Электронный ресурс], Воронеж, 1999, - 76 стр. Режим доступа: http://computersbooks.net/books/web-programmirovanie/zaencev-iv/1999/files/neyronnieseti1999.pdf

8. Материально-техническое обеспечение дисциплины.

Компьютерный класс Ауд. АВ4810, оснащенный методическими материалами по дисциплине (практические задания).

9. Методические рекомендации для самостоятельной работы студентов

Раздел дисциплины (тема)	Вид самостоятельной работы	Форма текущего контроля
Область применения хранилищ данных и больших данных и их преимущества	Самостоятельное изучение Изучение теоретического материала и решение задач	Письменное тестирование
Виды данных	Самостоятельное изучение Изучение теоретического материала и решение задач	Письменное тестирование
Экосистема больших данных	Самостоятельное изучение Изучение теоретического материала и решение задач	Письменное тестирование
Задачи Data Mining	Самостоятельное изучение Изучение теоретического материала и решение задач	Письменное тестирование
Задача кластеризации и EM-алгоритм	Самостоятельное изучение Изучение теоретического материала и решение задач	Письменное тестирование
Различные алгоритмы кластеризации	Самостоятельное изучение Изучение теоретического материала и решение задач	Письменное тестирование
Задача классификации	Самостоятельное изучение Изучение теоретического материала и решение задач	Письменное тестирование
Обработка текстов, Naive Bayes	Самостоятельное изучение Изучение теоретического материала и решение задач	Письменное тестирование
Линейные модели для классификации и регрессии	Самостоятельное изучение Изучение теоретического материала и решение задач	Письменное тестирование

Раздел дисциплины (тема)	Вид самостоятельной работы	Форма текущего контроля
Машина опорных векторов	Самостоятельное изучение Изучение теоретического материала и решение задач	Письменное тестирование
Методы снижения размерности пространства	Самостоятельное изучение Изучение теоретического материала и решение задач	Письменное тестирование
Алгоритмические композиции	Самостоятельное изучение Изучение теоретического материала и решение задач	Письменное тестирование
Глубокие нейронные сети	Самостоятельное изучение Изучение теоретического материала и решение задач	Письменное тестирование

10. Методические рекомендации для преподавателя

Тема занятий	Виды учебных занятий	Средства обучения	Методы обучения	Форма оценочного средства**
Область применения хранилищ данных и больших данных и их преимущества	Практическое занятие	Компьютерный класс	Практическое занятие	К-З
Виды данных	Практическое занятие	Компьютерный класс	Практическое занятие	К
Экосистема больших данных	Практическое занятие	Компьютерный класс	Практическое занятие	К-З
Задачи Data Mining	Практическое занятие	Компьютерный класс	Практическое занятие	К-З
Задача кластеризации и EM-алгоритм	Практическое занятие	Компьютерный класс	Практическое занятие	К-З
Различные алгоритмы кластеризации	Практическое занятие	Компьютерный класс	Практическое занятие	К-З
Задача классификации	Практическое занятие	Компьютерный класс	Практическое занятие	К-З
Обработка текстов, Naive Bayes	Практическое занятие	Компьютерный класс	Практическое занятие	К
Линейные модели для классификации и регрессии	Практическое занятие	Компьютерный класс	Практическое занятие	К-З
Машина опорных векторов	Практическое занятие	Компьютерный класс	Практическое занятие	К-З
Методы снижения размерности пространства	Практическое занятие	Компьютерный класс	Практическое занятие	К-З
Алгоритмические композиции	Практическое занятие	Компьютерный класс	Практическое занятие	К-З

Тема занятий	Виды учебных занятий	Средства обучения	Методы обучения	Форма оценочного средства**
Глубокие нейронные сети	Практическое занятие	Компьютерный класс	Практическое занятие	К

Приложение

Структура и содержание дисциплины «Хранилище данных» по направлению подготовки
Направление подготовки
09.03.03 «Прикладная информатика»
Квалификация (степень) выпускника
Бакалавр

n/n	Раздел	Семестр	Неделя семестра	Виды учебной работы, включая самостоятельную работу студентов, и трудоемкость в часах					Виды самостоятельной работы студентов					Формы аттестации	
				Л	П/С	Лаб	СРС	КСР	К.Р.	К.П.	К-З	Реферат	К/р	Э	З
	Седьмой семестр														
1	Задачи Data Mining	7	1	4		12	18								
2	Задача кластеризации и EM-алгоритм	7	2	4		12	18								
3	Различные алгоритмы кластеризации	7	3	4		12	18								
4	Задача классификации	7	4	4		12	18								
5	Обработка текстов, Naive Bayes	7	5	4		10	18								
6	Линейные модели для классификации и регрессии	8	6	4		10	18								
7	Машина опорных векторов	8	7	3		10	18								

n/n	Раздел	Семестр	Неделя семестра	Виды учебной работы, включая самостоятельную работу студентов, и трудоемкость в часах					Виды самостоятельной работы студентов					Формы аттестации	
				Л	П/С	Лаб	СРС	КСР	К.Р.	К.П.	К-З	Реферат	К/р	Э	З
8	Методы снижения размерности пространства	8	8	3		10	18								
9	Алгоритмические композиции	8	9	3		10	18								
10	Глубокие нейронные сети	8	10	3		10	18								
	Форма аттестации	8	11-12												З
	Всего часов по дисциплине в восьмом семестре			36		108	180								