

МИНИСТЕРСТВО ТРАНСПОРТА РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«РОССИЙСКИЙ УНИВЕРСИТЕТ ТРАНСПОРТА»
(РУТ (МИИТ))



Рабочая программа дисциплины (модуля),
как компонент образовательной программы
высшего образования - программы бакалавриата
по направлению подготовки
09.03.01 Информатика и вычислительная техника,
утвержденной первым проректором РУТ (МИИТ)
Тимониным В.С.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

Вводный курс "Специалист по Data Science" Яндекс Практикум

Направление подготовки: 09.03.01 Информатика и вычислительная техника

Направленность (профиль): IT-сервисы и технологии обработки данных на транспорте

Форма обучения: Очная

Рабочая программа дисциплины (модуля) в виде электронного документа выгружена из единой корпоративной информационной системы управления университетом и соответствует оригиналу

Простая электронная подпись, выданная РУТ (МИИТ)
ID подписи: 170737
Подписал: заместитель директора академии Паринов Денис Владимирович
Дата: 07.06.2024

1. Общие сведения о дисциплине (модуле).

Целью освоения учебной дисциплины является теоретическая и практическая подготовка студентов к работе с большими данными. Знания и компетенция, полученные в результате освоения дисциплины, помогут при сборе и анализе больших объемов структурированной или неструктурированной информации, при разработке моделей данных и получении новых знаний. Все это необходимо выпускнику, освоившему программу бакалавриата, для решения различных задач в области разработки корпоративных информационных систем и сервисов.

Задачи освоения дисциплины:

- приобретение студентами знаний о технологиях подготовки, хранения, обработки и анализа больших данных;
- применение статистических и математических методов для анализа больших объемов информации;

2. Планируемые результаты обучения по дисциплине (модулю).

Перечень формируемых результатов освоения образовательной программы (компетенций) в результате обучения по дисциплине (модулю):

ОПК-2 - Способен понимать принципы работы современных информационных технологий и программных средств, в том числе отечественного производства, и использовать их при решении задач профессиональной деятельности;

ПК-1 - Способен анализировать большие данные с использованием существующей в организации методологической и технологической инфраструктуры;

УК-1 - Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач.

Обучение по дисциплине (модулю) предполагает, что по его результатам обучающийся будет:

Знать:

- методы анализа, хранения и обработки больших данных, интегрируемых в корпоративные информационные системы,
- языки программирования, используемые для анализа больших данных,
- программное обеспечение, используемое для анализа больших данных

Уметь:

- использовать пакеты прикладных программ статистической обработки и анализа больших данных,
- определить инструменты, необходимые для статистического анализа больших данных исходя из целей анализа,
- разрабатывать независимые модули анализа данных.

Владеть:

- навыками сбора, загрузки и сохранения больших данных,
- навыками обучения регрессионных моделей
- навыками кластеризации данных

3. Объем дисциплины (модуля).

3.1. Общая трудоемкость дисциплины (модуля).

Общая трудоемкость дисциплины (модуля) составляет 2 з.е. (72 академических часа(ов)).

3.2. Объем дисциплины (модуля) в форме контактной работы обучающихся с педагогическими работниками и (или) лицами, привлекаемыми к реализации образовательной программы на иных условиях, при проведении учебных занятий:

Тип учебных занятий	Количество часов	
	Всего	Семестр №5
Контактная работа при проведении учебных занятий (всего):	6	6
В том числе:		
Занятия семинарского типа	6	6

3.3. Объем дисциплины (модуля) в форме самостоятельной работы обучающихся, а также в форме контактной работы обучающихся с педагогическими работниками и (или) лицами, привлекаемыми к реализации образовательной программы на иных условиях, при проведении промежуточной аттестации составляет 66 академических часа (ов).

3.4. При обучении по индивидуальному учебному плану, в том числе при ускоренном обучении, объем дисциплины (модуля) может быть реализован полностью в форме самостоятельной работы обучающихся, а также в форме контактной работы обучающихся с педагогическими работниками и (или) лицами, привлекаемыми к реализации образовательной программы на иных условиях, при проведении промежуточной аттестации.

4. Содержание дисциплины (модуля).

4.1. Занятия лекционного типа.

Не предусмотрено учебным планом

4.2. Занятия семинарского типа.

Практические занятия

№ п/п	Тематика практических занятий/краткое содержание
1	Тема 1. Регрессивный анализ данных. Рассматриваемые вопросы: - математическая постановка задачи регрессионного анализа - регрессионные модели - обучение регрессионных модели
2	Тема2. Кластерный анализ данных. Рассматриваемые вопросы: - методы кластеризации - методы поиска ближайшего соседа - анализ кластеров
3	Тема 3. Алгоритмы анализа больших данных Рассматриваемые вопросы: - математическая постановка задачи анализа больших данных - виды моделей машинного обучения

4.3. Самостоятельная работа обучающихся.

№ п/п	Вид самостоятельной работы
1	Работа с электронным курсом Яндекс
2	Подготовка к промежуточной аттестации.
3	Подготовка к промежуточной аттестации.

4.4. Примерный перечень тем курсовых работ

1

5. Перечень изданий, которые рекомендуется использовать при освоении дисциплины (модуля).

№ п/п	Библиографическое описание	Место доступа
1	Гудфеллоу, Я. Глубокое обучение / Я. Гудфеллоу, И. Бенджио, А. Курвилль ; перевод с английского А. А. Слинкина. — 2-е изд. — Москва : ДМК Пресс, 2018. — 652 с. — ISBN 978-5-97060-618-6	https://e.lanbook.com/book/107901

2	Габдуллин, Н. М. Развитие человеческого капитала и цифровой экономики в регионах России: факторный и кластерный анализ : монография / Н. М. Габдуллин. — Казань : КФУ, 2019. — 268 с. — ISBN 978-5-00130-291-9	https://e.lanbook.com/book/173018
3	Гласснер, Э. Глубокое обучение без математики. Том 2. Практика : руководство / Э. Гласснер ; перевод с английского В. А. Яроцкого. — Москва : ДМК Пресс, 2020. — 610 с. — ISBN 978-5-97060-767-1	https://e.lanbook.com/book/131710
4	Гулятьева, Т. А. Методы статистического обучения в задачах регрессии и классификации : монография / Т. А. Гулятьева, А. А. Попов, А. С. Саутин. — Новосибирск : НГТУ, 2016. — 323 с. — ISBN 978-5-7782-2817-7	https://e.lanbook.com/book/118291
5	Кук, Д. Машинное обучение с использованием библиотеки H2O / Д. Кук ; перевод с английского А. Б. Огурцова. — Москва : ДМК Пресс, 2018. — 250 с. — ISBN 978-5-97060-508-0	https://e.lanbook.com/book/97353
6	Флах, П. Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных / П. Флах. — Москва : ДМК Пресс, 2015. — 400 с. — ISBN 978-5-97060-273-7	https://e.lanbook.com/book/69955
7	Шалев-Шварц, Ш. Идеи машинного обучения : учебное пособие / Ш. Шалев-Шварц, Бен-Давид Ш. ; перевод с английского А. А. Слинкина. — Москва : ДМК Пресс, 2019. — 436 с. — ISBN 978-5-97060-673-5	https://e.lanbook.com/book/131686

6. Перечень современных профессиональных баз данных и информационных справочных систем, которые могут использоваться при освоении дисциплины (модуля).

<https://habr.com/ru> - база знаний в виде статей, обзоров

<https://journal.tinkoff.ru/short/ai-for-all/> - база данных нейронных сетей

<https://vc.ru/services/916617-luchshie-neyroseti-bolshaya-podborka-iz-top-200-ii-generatorov-po-kategoriyam> - база данных нейронных сетей

<https://github.com/abalmumcu/bert-rest-api> - профессиональная платформа для командой работы над проектов (нейронная сеть bert)

<http://library.miit.ru/> - электронно-библиотечная система Научно-технической библиотеки МИИТ

<https://proglib.io/p/raspoznavanie-obektov-s-pomoshchyu-yolo-v3-na-tensorflow-2-0-2020-11-08> - профессиональная библиотека программистов

https://yandex.cloud/ru/blog/posts/2022/12/andrey-berger-and-yandex-cloud?utm_referrer=https%3A%2F%2Fyandex.ru%2F – библиотека профессиональных статей разработчиков Яндекса

<https://yandex.cloud/ru/blog> - библиотека профессиональных статей разработчиков Яндекса

<https://tproger.ru/translations/opencv-python-guide> - библиотека основных команд OpenCV

7. Перечень лицензионного и свободно распространяемого программного обеспечения, в том числе отечественного производства, необходимого для освоения дисциплины (модуля).

1 Операционная система семейства Microsoft Windows

2 Пакет офисных программ Microsoft Office.

8. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине (модулю).

Компьютер преподавателя

Компьютеры студентов

экран для проектора, маркерная доска,

Проектор

9. Форма промежуточной аттестации:

Зачет в 5 семестре.

10. Оценочные материалы.

Оценочные материалы, применяемые при проведении промежуточной аттестации, разрабатываются в соответствии с локальным нормативным актом РУТ (МИИТ).

Авторы:

доцент, к.н. Академии "Высшая
инженерная школа"

Б.В. Игольников

Согласовано:

Заместитель директора академии

Д.В. Паринов

Председатель учебно-методической
комиссии

Д.В. Паринов