

**Федеральное государственное автономное образовательное  
учреждение высшего образования  
«Московский физико-технический институт  
(национальный исследовательский университет)»**

**УТВЕРЖДЕНО**

**Директор физтех-школы  
прикладной математики и  
информатики  
А.М. Райгородский**

|                            |                                                                            |
|----------------------------|----------------------------------------------------------------------------|
|                            | <b>Рабочая программа дисциплины (модуля)</b>                               |
| <b>по дисциплине:</b>      | Практическая аналитика                                                     |
| <b>по направлению:</b>     | Информатика и вычислительная техника                                       |
| <b>профиль подготовки:</b> | Физтех-школа Прикладной Математики и Информатики<br>кафедра анализа данных |
| <b>курс:</b>               | 4                                                                          |
| <b>квалификация:</b>       | бакалавр                                                                   |

Семестр, формы промежуточной аттестации: 7 (осенний) - Дифференцированный зачет

Аудиторных часов: 30 всего, в том числе:

лекции: 15 час.

семинары: 15 час.

лабораторные занятия: 0 час.

Самостоятельная работа: 15 час.

Всего часов: 45, всего зач. ед.: 1

Количество контрольных работ, заданий: 1

Программу составил: С.Н. Федотов, канд. физ.-мат. наук

Программа обсуждена на заседании кафедры анализа данных 04.06.2020

## Аннотация

Этот курс призван дать студентам опыт решения реальных практических задач анализа данных, а также глубже познакомить с теми разделами анализа данных и машинного обучения, которые важны при работе во многих IT-компаниях, но недостаточно хорошо покрыты в стандартных университетских курсах. Основная составляющая курса - самостоятельное решение задач.

### 1. Цели и задачи

#### Цель дисциплины

- подготовить студентов к решению индустриальных и бизнес-задач с использованием анализа данных.

#### Задачи дисциплины

- познакомить студентов с этапами решения задач анализа данных, научить выбирать и измерять метрики; дать опыт решения практически задач.

### 2. Перечень формируемых компетенций

Освоение дисциплины направлено на формирование следующих компетенций:

| Код и наименование компетенции                                                                                                                                                  | Индикаторы достижения компетенции                                                                                                                                                 |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| УК-1 Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач                                             | УК-1.1 Анализирует задачу, выделяя этапы ее решения, действия по решению задачи                                                                                                   |
|                                                                                                                                                                                 | УК-1.2 Находит, критически анализирует и выбирает информацию, необходимую для решения поставленной задачи                                                                         |
|                                                                                                                                                                                 | УК-1.3 Рассматривает различные варианты решения задачи, оценивает их преимущества и недостатки                                                                                    |
|                                                                                                                                                                                 | УК-1.4 Грамотно, логично, аргументированно формирует собственные суждения и оценки                                                                                                |
|                                                                                                                                                                                 | УК-1.5 Определяет и оценивает практические последствия возможных вариантов решения задачи                                                                                         |
| УК-2 Способен определять круг задач в рамках поставленной цели и выбирать оптимальные способы их решения, исходя из действующих правовых норм, имеющихся ресурсов и ограничений | УК-2.1 Формулирует совокупность взаимосвязанных задач в рамках поставленной цели работы, обеспечивающих ее достижение. Определяет ожидаемые результаты решения поставленных задач |
|                                                                                                                                                                                 | УК-2.2 Проектирует решение конкретной задачи проекта, выбирая оптимальный способ ее решения, исходя из действующих правовых норм и имеющихся ресурсов и ограничений               |
| ОПК-4 Способен осуществлять сбор и обработку научно-технической и (или) технологической информации для решения фундаментальных и прикладных задач                               | ОПК-4.1 Владеет методами научного поиска и интеллектуального анализа информации при решении задач профессиональной деятельности                                                   |
|                                                                                                                                                                                 | ОПК-4.2 Знает основные источники научно-технической и (или) технологической информации в области профессиональной деятельности                                                    |
|                                                                                                                                                                                 | ОПК-4.3 Умеет составлять аннотации, рефераты, библиографические перечни и обзоры информации в области своей профессиональной деятельности                                         |
|                                                                                                                                                                                 | ОПК-4.4 Владеет навыками работы с компьютером и компьютерными сетями с целью получения, хранения и обработки научной (технической, технологической) информации                    |
| ОПК-5 Способен участвовать в проведении фундаментальных и прикладных исследований и разработок, самостоятельно осваивать новые теоретические, в том числе,                      | ОПК-5.3 Способен к профессиональной эксплуатации современной экспериментальной научно-исследовательской (измерительно-аналитической и технологической) аппаратуры                 |

|                                                                                                                                                                                                                                 |                                                                                                                                                                        |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| математические методы исследований и работать на современной экспериментальной научно-исследовательской, измерительно-аналитической и технологической аппаратуре)                                                               | ОПК-5.1 Способен решать поставленные задачи в области теоретических и экспериментальных исследований и разработок                                                      |
|                                                                                                                                                                                                                                 | ОПК-5.2 Обладает способностью к освоению новых знаний на основе изучения литературы, научных статей и других источников                                                |
| ПК-1 Способен ставить, формализовывать и решать задачи, в том числе разрабатывать и исследовать математические модели изучаемых явлений и процессов, системно анализировать научные проблемы, получать новые научные результаты | ПК-1.1 Способен находить, анализировать и обобщать информацию об актуальных результатах исследований в рамках тематической области своей профессиональной деятельности |
|                                                                                                                                                                                                                                 | ПК-1.2 Способен выдвигать гипотезы, строить математические модели для описания изучаемых явлений и процессов, оценить качество разработанной модели                    |
|                                                                                                                                                                                                                                 | ПК-1.3 Способен применять теоретические и (или) экспериментальные методы исследований к конкретной научной задаче и интерпретировать полученные результаты             |
| ПК-2 Способен самостоятельно или в качестве члена (руководителя) малого коллектива организовывать и проводить научные исследования и их апробацию                                                                               | ПК-2.1 Знает принципы построения научной работы, методы сбора и анализа полученного материала, способы аргументации                                                    |
|                                                                                                                                                                                                                                 | ПК-2.2 Способен планировать и проводить научные исследования самостоятельно или в качестве члена (руководителя) малого научного коллектива                             |
|                                                                                                                                                                                                                                 | ПК-2.3 Способен проводить апробацию результатов научно-исследовательской работы посредством публикации научных статей и участия в конференциях                         |

### 3. Перечень планируемых результатов обучения по дисциплине (модулю)

В результате освоения дисциплины обучающиеся должны

знать:

- как понять, будет ли то или иное решение полезно для бизнеса.

уметь:

- выстраивать процесс решения задачи анализа данных, получать данные и выбирать метрики.

владеть:

- основными принципами А/Б-тестирования.

### 4. Содержание дисциплины (модуля), структурированное по темам (разделам) с указанием отведенного на них количества академических часов и видов учебных занятий

#### 4.1. Разделы дисциплины (модуля) и трудоемкости по видам учебных занятий

| № | Тема (раздел) дисциплины                        | Трудоемкость по видам учебных занятий, включая самостоятельную работу, час. |          |                 |                |
|---|-------------------------------------------------|-----------------------------------------------------------------------------|----------|-----------------|----------------|
|   |                                                 | Лекции                                                                      | Семинары | Лаборат. работы | Самост. работа |
| 1 | Пайплайн решения задач анализа данных.          | 2                                                                           | 2        |                 | 2              |
| 2 | Предсказание временных рядов.                   | 2                                                                           | 2        |                 | 2              |
| 3 | Интерпретируемость моделей и этические вопросы. | 2                                                                           | 2        |                 | 2              |
| 4 | Критерии качества в задачах анализа данных.     | 1                                                                           | 1        |                 | 1              |
| 5 | Онлайн-метрики и А/Б-тестирование.              | 1                                                                           | 1        |                 | 1              |
| 6 | Кластеризация.                                  | 2                                                                           | 2        |                 | 2              |

|                       |                                               |                    |    |  |    |
|-----------------------|-----------------------------------------------|--------------------|----|--|----|
| 7                     | Предсказание позиций в поисковой выдаче.      | 1                  | 1  |  | 1  |
| 8                     | Предсказание кликов.                          | 1                  | 1  |  | 1  |
| 9                     | Свежесть рекомендаций.                        | 1                  | 1  |  | 1  |
| 10                    | Итоги: чем Data Science может помочь бизнесу? | 2                  | 2  |  | 2  |
| Итого часов           |                                               | 15                 | 15 |  | 15 |
| Подготовка к экзамену |                                               | 0 час.             |    |  |    |
| Общая трудоёмкость    |                                               | 45 час., 1 зач.ед. |    |  |    |

#### 4.2. Содержание дисциплины (модуля), структурированное по темам (разделам)

Семестр: 7 (Осенний)

##### 1. Пайплайн решения задач анализа данных.

Этапы решения задачи анализа данных. Командный разбор кейсов.

##### 2. Предсказание временных рядов.

Модели для предсказания рядов. Оценка качества моделей.

##### 3. Интерпретируемость моделей и этические вопросы.

Что такое интерпретируемость и зачем она нужна. Интерпретируемые методы и model-agnostic методы интерпретации. Этические вопросы в задачах анализа данных.

##### 4. Критерии качества в задачах анализа данных.

Метрики в задачах анализа данных. Продуктовые и бизнес-метрики. Подбор метрик в ситуациях, когда критерии успеха не диктуются самой задачей.

##### 5. Онлайн-метрики и А/Б-тестирование.

Онлайн-оценка моделей. Эксперименты и А/Б-тестирование. Планирование экспериментов. Статистическая оценка результатов экспериментов.

##### 6. Кластеризация.

Методы кластеризации. Кластеризация для очень больших датасетов.

##### 7. Предсказание позиций в поисковой выдаче.

Разбор кейса: предсказание позиции колдунщика в поисковой выдаче.

##### 8. Предсказание кликов.

Разбор кейса: построение оффлайн-метрики числа кликов на выдаче поиска.

##### 9. Свежесть рекомендаций.

Что такое свежесть и зачем она нужна. Разбор кейса про свежесть рекомендаций.

##### 10. Итоги: чем Data Science может помочь бизнесу?

Командный разбор кейсов.

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

учебная аудитория, оснащенная компьютером и мультимедийным оборудованием (проектор, звуковая система).

## **6.Перечень рекомендуемой литературы**

### Основная литература

1. Методы анализа данных [Текст] = Optimisation en classification automatique, подход, основанный на методе динамических сгущений/пер. с фр. С. А. Айвазяна, В. Д. Конакова, С. Ю. Адамова, -М., Финансы и статистика, 1985

### Дополнительная литература

1. Основы Data Science и Big Data. Python и наука о данных [Текст]/Д. Силен, А. Мейсман, М. Али, -СПб., Питер, 2017

## **7. Перечень ресурсов информационно-телекоммуникационной сети "Интернет", необходимых для освоения дисциплины (модуля)**

<https://web.stanford.edu/~hastie/ElemStatLearn/>  
<https://www.deeplearningbook.org/>

## **8. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень необходимого программного обеспечения и информационных справочных систем (при необходимости)**

на занятиях используются мультимедийные технологии, включая демонстрацию презентаций.

## **9. Методические указания для обучающихся по освоению дисциплины (модуля)**

Методические рекомендации позволяют студенту оптимальным образом организовать процесс обучения. В рабочей программе приведено примерное распределение часов аудиторной и внеаудиторной нагрузки по различным темам данной дисциплины.

**ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)**

**по направлению:** Информатика и вычислительная техника

**профиль подготовки:** Физтех-школа Прикладной Математики и Информатики  
кафедра анализа данных

**курс:** 4

**квалификация:** бакалавр

Семестр, формы промежуточной аттестации: 7 (осенний) - Дифференцированный зачет

**Разработчик:** С.Н. Федотов, канд. физ.-мат. наук

# 1. Компетенции, формируемые в процессе изучения дисциплины

| Код и наименование компетенции                                                                                                                                                                                                                                                                                               | Индикаторы достижения компетенции                                                                                                                                                 |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| УК-1 Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач                                                                                                                                                                                          | УК-1.1 Анализирует задачу, выделяя этапы ее решения, действия по решению задачи                                                                                                   |
|                                                                                                                                                                                                                                                                                                                              | УК-1.2 Находит, критически анализирует и выбирает информацию, необходимую для решения поставленной задачи                                                                         |
|                                                                                                                                                                                                                                                                                                                              | УК-1.3 Рассматривает различные варианты решения задачи, оценивает их преимущества и недостатки                                                                                    |
|                                                                                                                                                                                                                                                                                                                              | УК-1.4 Грамотно, логично, аргументированно формирует собственные суждения и оценки                                                                                                |
|                                                                                                                                                                                                                                                                                                                              | УК-1.5 Определяет и оценивает практические последствия возможных вариантов решения задачи                                                                                         |
| УК-2 Способен определять круг задач в рамках поставленной цели и выбирать оптимальные способы их решения, исходя из действующих правовых норм, имеющихся ресурсов и ограничений                                                                                                                                              | УК-2.1 Формулирует совокупность взаимосвязанных задач в рамках поставленной цели работы, обеспечивающих ее достижение. Определяет ожидаемые результаты решения поставленных задач |
|                                                                                                                                                                                                                                                                                                                              | УК-2.2 Проектирует решение конкретной задачи проекта, выбирая оптимальный способ ее решения, исходя из действующих правовых норм и имеющихся ресурсов и ограничений               |
| ОПК-4 Способен осуществлять сбор и обработку научно-технической и (или) технологической информации для решения фундаментальных и прикладных задач                                                                                                                                                                            | ОПК-4.1 Владеет методами научного поиска и интеллектуального анализа информации при решении задач профессиональной деятельности                                                   |
|                                                                                                                                                                                                                                                                                                                              | ОПК-4.2 Знает основные источники научно-технической и (или) технологической информации в области профессиональной деятельности                                                    |
|                                                                                                                                                                                                                                                                                                                              | ОПК-4.3 Умеет составлять аннотации, рефераты, библиографические перечни и обзоры информации в области своей профессиональной деятельности                                         |
|                                                                                                                                                                                                                                                                                                                              | ОПК-4.4 Владеет навыками работы с компьютером и компьютерными сетями с целью получения, хранения и обработки научной (технической, технологической) информации                    |
| ОПК-5 Способен участвовать в проведении фундаментальных и прикладных исследований и разработок, самостоятельно осваивать новые теоретические, в том числе, математические методы исследований и работать на современной экспериментальной научно-исследовательской, измерительно-аналитической и технологической аппаратуре) | ОПК-5.3 Способен к профессиональной эксплуатации современной экспериментальной научно-исследовательской (измерительно-аналитической и технологической) аппаратуры                 |
|                                                                                                                                                                                                                                                                                                                              | ОПК-5.1 Способен решать поставленные задачи в области теоретических и экспериментальных исследований и разработок                                                                 |
|                                                                                                                                                                                                                                                                                                                              | ОПК-5.2 Обладает способностью к освоению новых знаний на основе изучения литературы, научных статей и других источников                                                           |
| ПК-1 Способен ставить, формализовывать и решать задачи, в том числе разрабатывать и исследовать математические модели изучаемых явлений и процессов, системно анализировать научные проблемы, получать новые научные результаты                                                                                              | ПК-1.1 Способен находить, анализировать и обобщать информацию об актуальных результатах исследований в рамках тематической области своей профессиональной деятельности            |
|                                                                                                                                                                                                                                                                                                                              | ПК-1.2 Способен выдвигать гипотезы, строить математические модели для описания изучаемых явлений и процессов, оценить качество разработанной модели                               |
|                                                                                                                                                                                                                                                                                                                              | ПК-1.3 Способен применять теоретические и (или) экспериментальные методы исследований к конкретной научной задаче и интерпретировать полученные результаты                        |

|                                                                                                                                                   |                                                                                                                                                |
|---------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------|
| ПК-2 Способен самостоятельно или в качестве члена (руководителя) малого коллектива организовывать и проводить научные исследования и их апробацию | ПК-2.1 Знает принципы построения научной работы, методы сбора и анализа полученного материала, способы аргументации                            |
|                                                                                                                                                   | ПК-2.2 Способен планировать и проводить научные исследования самостоятельно или в качестве члена (руководителя) малого научного коллектива     |
|                                                                                                                                                   | ПК-2.3 Способен проводить апробацию результатов научно-исследовательской работы посредством публикации научных статей и участия в конференциях |

## 2. Показатели оценивания компетенций

В результате изучения дисциплины «Практическая аналитика» обучающийся должен:

### знать:

- как понять, будет ли то или иное решение полезно для бизнеса.

### уметь:

- выстраивать процесс решения задачи анализа данных, получать данные и выбирать метрики.

### владеть:

- основными принципами А/Б-тестирования.

## 3. Перечень типовых (примерных) вопросов, заданий, тем для подготовки к текущему контролю

1. Этапы решения задачи анализа данных.
2. Командный разбор кейсов.
3. Модели для предсказания рядов.
4. Оценка качества моделей.
5. Что такое интерпретируемость и зачем она нужна.
6. Интерпретируемые методы и model-agnostic методы интерпретации.
7. Этические вопросы в задачах анализа данных.
8. Метрики в задачах анализа данных.
9. Продуктовые и бизнес-метрики.
10. Подбор метрик в ситуациях, когда критерии успеха не диктуются самой задачей.
11. Онлайн-оценка моделей.
12. Эксперименты и А/Б-тестирование.
13. Планирование экспериментов.
14. Статистическая оценка результатов экспериментов.
15. Методы кластеризации.
16. Кластеризация для очень больших датасетов.
17. Разбор кейса: предсказание позиции колдунщика в поисковой выдаче.
18. Разбор кейса: построение оффлайн-метрики числа кликов на выдаче поиска.
19. Что такое свежесть и зачем она нужна.
20. Разбор кейса про свежесть рекомендаций.
21. Командный разбор кейсов.

## 4. Перечень типовых (примерных) вопросов и тем для проведения промежуточной аттестации обучающихся

1. Вас просят проверить гипотезу о том, влияют ли погодные условия на пробки. Объясните, как вы стали бы это делать.
2. Требуется определять "плохие" вопросы в вопросно-ответной системе. Опишите возможный пайплайн решения задачи.
3. Опишите метод интерпретации моделей LIME. Что он позволяет сделать? Каковы его достоинства и недостатки?
4. Требуется исправить модель классификации так, чтобы она не учитывала пол человека. Как это можно сделать?



5. Опишите какую-нибудь архитектуру рекомендательной системы. Как её обучать? Что делать с новыми пользователями и с новыми объектами системы?

#### Критерии оценивания

отлично (10) - выставляется студенту, показавшему всесторонние, систематизированные, глубокие знания учебной программы дисциплины и умение уверенно применять их на практике при решении конкретных задач, свободное и правильное обоснование принятых решений.

отлично (9) - выставляется студенту, показавшему свободное оперирование знаниями учебной программы дисциплины, выполнение заданий творческого характера.

отлично (8) - выставляется студенту, показавшему владение программным учебным материалом с наличием несущественных ошибок в действиях, самостоятельно исправляемых учащимся.

хорошо (7) - выставляется студенту, если он твердо знает материал, грамотно и по существу излагает его, умеет применять полученные знания на практике, но допускается в ответе или в решении задач некоторые неточности.

хорошо (6) - выставляется студенту если он осознает воспроизведение программного учебного материала, в том числе и различной степени сложности, с несущественными ошибками, затруднения в применении отдельных навыков.

хорошо (5) - выставляется студенту если теоретическое содержание освоено не полностью, некоторые практические навыки сформированы недостаточно, в некоторых случаях были допущены ошибки.

удовлетворительно (4) - выставляется студенту, показавшему фрагментарный, разрозненный характер знаний, недостаточно правильные формулировки базовых понятий, нарушения логической последовательности в изложении программного материала, но при этом он владеет основными разделами учебной программы, необходимыми для дальнейшего обучения и может применять полученные знания по образцу в стандартной ситуации.

удовлетворительно (3) - выставляется студенту в случае большого количества недочетов и неправильных ответов, а также пассивной работе в ходе занятий, многие учебные задания не выполнены.

неудовлетворительно (2) - выставляется студенту, который не знает большей части основного содержания учебной программы дисциплины, допускает грубые ошибки в формулировках основных понятий дисциплины и не умеет использовать полученные знания при решении типовых практических задач.

неудовлетворительно (1) - выставляется студенту, который не освоил теоретическое и практическое содержание курса, все выполненные учебные задания содержат грубые ошибки.

#### **5. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности**

При проведении дифференцированного зачета обучающемуся предоставляется 30 минут на подготовку. Опрос обучающегося по билету на дифференцированном зачете не должен превышать одного астрономического часа.

Во время проведения дифференцированного зачета обучающиеся могут пользоваться программой дисциплины, а также справочной литературой, конспектами лекций или другими материалами.