



Национальный исследовательский университет «Высшая школа экономики»  
Программа дисциплины «Управление ИТ-проектами и командами разработчиков в сфере data  
Science» для направления 01.04.02 Образовательной программы «Прикладная математика и ин-  
форматика»  
подготовки магистра

**Федеральное государственное автономное образовательное учреждение  
высшего образования  
«Национальный исследовательский университет  
«Высшая школа экономики»**

Факультет компьютерных наук  
Департамент больших данных и информационного поиска

**Рабочая программа дисциплины**  
Управление ИТ-проектами и командами разработчиков в сфере Data Science

для образовательной программы «Прикладная математика и информатика»  
направления 01.04.02  
уровень магистр

Разработчик программы  
Белугин А.А., thedragon@yandex-team.ru

Одобрена на заседании департамента больших данных и информационного поиска  
«\_\_»\_\_\_\_\_ 2017г.

Руководитель департамента  
В.В.Подольский \_\_\_\_\_ [подпись]

Утверждена Академическим советом образовательной программы  
«\_\_»\_\_\_\_\_ 2017 г., № протокола \_\_\_\_\_

Академический руководитель образовательной программы  
А.А. Масютин \_\_\_\_\_ [подпись]

Москва, 2017

*Настоящая программа не может быть использована другими подразделениями университета и  
другими вузами без разрешения подразделения-разработчика программы.*



## 1 Область применения и нормативные ссылки

Настоящая программа учебной дисциплины устанавливает требования к образовательным результатам и результатам обучения студента и определяет содержание и виды учебных занятий и отчетности.

Программа предназначена для преподавателей, ведущих дисциплину Управление ИТ-проектами и командами разработчиков в сфере Data Science, учебных ассистентов и студентов направления подготовки 01.04.02 «Прикладная математика и информатика», обучающихся по образовательной программе «Финансовые технологии и анализ данных».

Программа учебной дисциплины разработана в соответствии с:

- образовательным стандартом Федерального государственного автономного образовательного учреждения высшего образования Национальный исследовательский университет «Высшая школа экономики», уровень подготовки: магистр, утвержденным ученым советом НИУ ВШЭ 22.12.2017 протокол №13;
- рабочим учебным планом университета по направлению 01.04.02 «Прикладная математика и информатика» подготовки магистра для магистерской программы «Финансовые технологии и анализ данных», утвержденным 02.02.2018 г.

## 2 Цели освоения дисциплины

Целями освоения дисциплины «Управление ИТ-проектами и командами разработчиков в сфере Data Science» являются:

- Ознакомление студентов с теоретическими основами и основными принципами управления проектами с применением технологий машинного обучения, оценки объема и сроков работ.
- Формирование у студентов практических навыков постановки задач машинного обучения и оценки результатов.

## 3 Компетенции обучающегося, формируемые в результате освоения дисциплины

В результате освоения дисциплины студент осваивает компетенции:

Код по ОС ВШЭ	Уровень формирования компетенции	Дескрипторы – основные признаки освоения (показатели достижения результата)	Формы и методы обучения, способствующие формированию и развитию компетенции	Форма контроля уровня сформированности компетенции
ПК-1	РБ СД	Способен описывать проблемы и ситуации профессиональной деятельности, используя язык и аппарат математики	Аудиторная работа на семинарах	Коллоквиум, экзамен
К5	СК-М5	Способен принимать управленческие решения и готов нести за них ответственность.	Аудиторная работа на семинарах, выполнение теоретических домашних заданий	Домашние задания



Код по ОС ВШЭ	Уровень формирования компетенции	Дескрипторы – основные признаки освоения (показатели достижения результата)	Формы и методы обучения, способствующие формированию и развитию компетенции	Форма контроля уровня сформированности компетенции
К6	СК-М6	Способен анализировать, верифицировать, оценивать полноту информации в ходе профессиональной деятельности, при необходимости восполнять и синтезировать недостающую информацию.	Выполнение практических домашних заданий	Домашние задания
К7	СК-М7	Способен организовать многостороннюю коммуникацию и управлять ею.	Выполнение практических домашних заданий	Домашние задания
К8	СК-М8	Способен вести профессиональную, в том числе научно-исследовательскую деятельность в международной среде	Выполнение практических домашних заданий	Домашние задания

#### 4 Место дисциплины в структуре образовательной программы

Настоящая дисциплина относится к циклу дисциплин по машинному обучению и анализу данных.

Для освоения учебной дисциплины студенты должны владеть знаниями и компетенциями по:

- математическому анализу, линейной алгебре и теории вероятностей, которые проверяются в рамках вступительных экзаменов на магистерскую программу;
- основам машинного обучения, преподаваемым на 1 курсе магистерской программы.

Основные положения дисциплины должны быть использованы в дальнейшем при изучении дисциплин:

- Современные методы принятия решений: алгоритмы обработки больших данных
- Управление данными и исполнение моделей

#### 5 Тематический план учебной дисциплины

№	Название раздела	Всего часов	Аудиторные часы		Самостоятельная работа
			Лекции	Семинары	
1	Обзор объектов управления	48	4	4	40
2	Постановка задачи	48	4	4	40
3	Оценка качества	48	4	4	40
4	Планирование проекта	36	4	4	38



№	Название раздела	Всего часов	Аудиторные часы		Самостоятельная работа
	ИТОГО	190	16	16	158

## 6 Формы контроля знаний студентов

Тип контроля	Форма контроля	1 год	Параметры
		2	
Текущий	Домашнее задание	*	Практическая работа: формальная постановка задачи машинного обучения
	Домашнее задание	*	Практическая работа: оценка качества предсказательной модели
	Домашнее задание	*	Практическая работа: подготовка плана проекта
Итоговый	Экзамен	*	Письменный

## 7 Критерии оценки знаний, навыков

В курсе предусмотрено несколько форм контроля знания:

- Самостоятельные работы на семинарах, проверяющие знание основных фактов с лекций и выполнение практических домашних заданий
- Практические домашние работы, формирующие навыки работы с проектами машинного обучения
- Письменный экзамен

Оценки по всем формам текущего контроля выставляются по 10-ти балльной шкале.

## 8 Содержание дисциплины

### 8.1 Обзор объектов управления

Лекция №1. Этапы проекта. Разбор примера плана проекта с указанием ролей, задач, проблем. Рынок данных. Юридические аспекты работы с данными.

Лекция №2. Инструменты аналитика — python, R, ноутбуки; SAS, SPSS, RapidMiner; библиотеки — какие и зачем; Hadoop, Spark.

Семинары №1 и №2 — пример решения задач на python и на SPSS.

### 8.2 Постановка задачи

Лекция №3. Типы задач, примеры; выбор метрик - вводная часть; управление ожиданиями по качеству; анализ сценариев использования моделей.

Лекция №4. Валидация задачи; валидация данных; экономический эффект.

Семинары №3 и №4 — примеры задач, неправильных постановок, работа с данными с помощью SPSS.

### 8.3 Оценка качества

Лекция №5. Метрики классификации; метрики регрессии; метрики ранжирования; стоимостное выражение метрик.



Лекция №6. Как и зачем мерять качество; А/Б тестирование; стат.значимость, чувствительность; примеры онлайн и офлайн тестов.

Семинары №5 и №6 — работа с данными результатов различных тестов и их оценка с помощью SPSS.

#### **8.4 Планирование проекта**

Лекция №7. Состав работ в проекте на детальном уровне; предобработка данных; планирование экспериментов; валидация и тестирование в ходе проекта

Лекция №8. Разбор примеров проектов с учетом всего ранее изученного

Семинары №7 и №8 — разбор примеров задач, работа с данными с помощью SPSS.

### **9 Образовательные технологии**

Необходимое для выполнения домашних заданий программное обеспечение находится в свободном доступе и его можно загрузить в сети Интернет. К каждой теме студентам выдаются конспекты или слайды лекций и семинаров, а также программный код, иллюстрирующий практическое использование изучаемых методов.

На практических семинарах используются программные продукты Microsoft Office Excel и IBM SPSS Modeler. Программами можно воспользоваться в учебных классах.

### **10 Оценочные средства для текущего контроля и аттестации студента**

#### **10.1 Оценочные средства для оценки качества освоения дисциплины в ходе текущего контроля**

Структура домашних заданий (практических): дано описание бизнес-процесса, сведения о наличии тех или иных данных. Необходимо:

1. Подготовить постановку задачи с указанием целей, метрики и порядка оценки проекта по машинному обучению
2. Рассмотреть различные варианты тестирования и оценки качества, сделать выводы
3. Подготовить план проекта по применению технологии машинного обучения

#### **10.2 Примеры заданий итоговой аттестации**

Примеры экзаменационных вопросов:

1. Назовите этапы проекта по машинному обучению
2. Перечислите основные группы инструментов специалистов по машинному обучению, сферы их применения
3. Какие законы регулируют работу с данными? Опишите основные моменты
4. Перечислите шаги валидации задачи машинного обучения
5. Перечислите шаги валидации данных для задач машинного обучения
6. Перечислите основные методы оценки качества моделей при промышленном внедрении

### **11 Порядок формирования оценок по дисциплине**

Результирующая оценка по дисциплине рассчитывается по формуле:

$$O_{\text{итог}} = 0,6 O_{\text{накопл}} + 0,4 O_{\text{экз}}$$

Накопленная оценка рассчитывается по формуле:



$$O_{\text{накопл}} = 0,6 O_{\text{ДЗ}} + 0,4 O_{\text{семинар}}$$

Накопленная и итоговая оценки округляются арифметически.

Оценка за домашние задания рассчитывается как среднее значение оценок за все выданные домашние задания (всего три задания). Домашнее задание имеет срок выполнения, каждые 7 дней просрочки вычитается 1 балл из максимально возможной оценки. Например, сдав задание через 15 дней после срока, максимальный балл составит 8.

Оценка за работу на семинаре оценивается как средняя оценка за все семинары. Таким образом, каждый семинар вносит 0,05 в накопленную оценку. Оценка за конкретный семинар выставляется по факту выполнения задания по итогам завершения курса:

- 0 — не выполнено;
- 6–7 — выполнено позже чем через 14 дней после семинара;
- 8–10 — выполнено на семинаре или в течение 14 дней после

Для примера, если семинар прошел 1 ноября, то 15 ноября в 23:59 последняя минута, когда можно получить 8–10 баллов за задание.

## 12 Учебно-методическое и информационное обеспечение дисциплины

### 12.1 Базовые учебники

1. David Julian, Designing Machine Learning Systems with Python, PACKT, 2016
2. Gene Kim, Kevin Behr, George Spafford, The Phoenix Project: A Novel About IT, DevOps, and Helping Your Business Win, IT Revolution Press, 2014
3. Jennifer Davis, Katherine Daniels, Effective DevOps: Building a Culture of Collaboration, Affinity, and Tooling at Scale, O'Reilly Media, Inc., 2016
4. Mark C. Layton, Agile Project Management For Dummies, John Wiley & Sons, 2012

### 12.2 Дополнительная литература

4. Как понять, что ваша предсказательная модель бесполезна; <https://habrahabr.ru/post/337722/>
5. Метрики в задачах машинного обучения, <https://habrahabr.ru/company/ods/blog/328372/>
6. Байесовские многорукие бандиты против A/B тестов, <https://habrahabr.ru/company/ods/blog/325416/>