

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

ВЫПИСКА ИЗ ПРОТОКОЛА № 10
заседания учебно-методического совета от 22 июня 2021 года.

ПОВЕСТКА:

Рассмотрение дополнительных общеобразовательных и профессиональных программ.

Проректор по учебной работе А. А. Воронов

СЛУШАЛИ: директора по внутреннему контролю и аудиту Е. Г. Евсева о представлении дополнительных общеобразовательных и профессиональных программ. (ЦДПО).

ПОСТАНОВИЛИ:

Рекомендовать к утверждению в установленном порядке дополнительную профессиональную программу повышения квалификации «Python и инструменты машинного обучения»

Решение принято единогласно заочным голосованием.

Форма проведения заседания: заочная

Председатель УМС МФТИ

Ученый секретарь УМС МФТИ



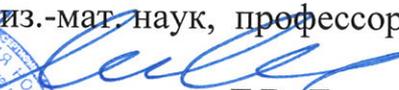
Министерство науки и высшего образования Российской Федерации
федеральное государственное автономное образовательное учреждение
высшего образования

«Московский физико-технический институт
(национальный исследовательский университет)»
(МФТИ)

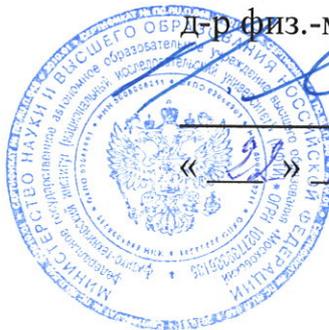
УТВЕРЖДАЮ

И.о. ректора МФТИ

д-р физ.-мат. наук, профессор


Д.В. Ливанов

«» 2021 г.



**Дополнительная профессиональная
программа повышения квалификации
«Python и инструменты машинного обучения»**

УГСН 02.00.00 Компьютерные и информационные науки
ФГОС 02.03.01 «Математика и компьютерные науки (уровень бакалавриата)»

ОКВЭД 70.10.1 «Деятельность по управлению финансово-промышленными группами»
ОКВЭД 63.11.1 «Деятельность по созданию и использованию баз данных и
информационных ресурсов»

Москва 2021

1. Общая характеристика программы

1.1 Цель реализации программы

Целью реализации дополнительной профессиональной программы повышения квалификации (далее - Программа) «Python и инструменты машинного обучения» является приобретение и совершенствование профессиональных, общепрофессиональных, универсальных и общекультурных компетенций слушателей в сфере прикладной математической статистики, анализа данных.

Совершенствуемые компетенции

Компетенции формируемые и совершенствуемые в результате обучения представлены в таблицах 1 и 2.

Таблица 1

№	Компетенция в соответствии с направлением подготовки	«Прикладная математика и информатика (уровень бакалавриата)» код: 01.03.02
1.	Способность демонстрации общенаучных базовых знаний естественных наук, математики и информатики, понимание основных фактов, концепций, принципов теорий, связанных с прикладной математикой и информатикой	ПК-1
2.	Способность понимать и применять в исследовательской и прикладной деятельности современный математический аппарат	ПК-2
3.	Способность критически переосмысливать накопленный опыт, изменять при необходимости вид и характер своей профессиональной деятельности	ПК-3
4.	Способность применять в профессиональной деятельности современные языки программирования, библиотеки и пакеты программ, а также уверенно владеть методами машинного обучения	ПК-4

Таблица 2

№	Универсальные компетенций (УК), общекультурные компетенций (ОК) и общепрофессиональные компетенций	«Прикладная математика и информатика (уровень бакалавриата)» код: 01.03.02
---	--	--

1.	Способность к критическому анализу и оценке современных научных достижений, генерированию новых идей при решении исследовательских и практических задач, в том числе в междисциплинарных областях	УК-1
2.	Способность к абстрактному мышлению, анализу, синтезу, способность совершенствовать и развивать свой интеллектуальный и общекультурный уровень	ОК-2
3.	Готовность действовать в нестандартных ситуациях, нести социальную и этическую ответственность за принятые решения	ОК-3
4.	Способность использовать базовые знания естественных наук, математики и информатики, основные факты, концепции, принципы теорий, связанных с прикладной математикой и информатикой в контексте анализа данных и смежных задач	ОПК-1
5	Способность приобретать новые научные и профессиональные знания, используя современные образовательные и информационные технологии	ОПК-2

1.2. Планируемые результаты обучения

Планируемые результаты обучения представлены в таблице 3.

Таблица 3

№	Знать	Коды компетенций
1.	Принцип работы алгоритмов классификации	ПК-1, ОПК-2
2.	Принцип работы алгоритмов регрессии	
3.	Принцип работы методов отбора признаков (одномерные, жадные, на основе моделей)	ПК-1, ОПК-2
4.	Постановка задачи обучения по прецедентам	ПК-1, ОПК-2
5.	Критерии выбора модели для построения прогноза	ПК-1, ОПК-2
6.	Язык Python и библиотеки для анализа данных (NumPy, SciPy, Pandas, Matplotlib, Sklearn)	ПК-1, ОПК-2
8.	Процедуру вычисления значения метрик качества моделей	ПК-1, ОПК-2
10.	Методы работы с естественным языком	ПК-1, ОПК-2
11.	Анализ временных рядов методами машинного обучения	ПК-1, ОПК-2
№	Уметь	Коды компетенций
1.	Решать задачу классификации данных	ПК-2, ОК-2
	Решать задачу регрессии на данных	ПК-2, ОК-2

2.	Выбрать алгоритм классификации или регрессии, оптимальный для конкретной задачи	ПК-2, ПК-3, УК-1
3.	Перечислить и охарактеризовать метрики качества	ПК-2, ОК-2, ОПК-1
4.	Применить средства sklearn для предобработки и процессинга объектов	ПК-2, ПК-3, ОПК-1
5.	Описать решение задачи классификации	ПК-2, ОК-2
6.	Описать решение задачи регрессии	ПК-2, ОК-2
7.	Применить на практике классификацию	ПК-2, ПК-3, УК-1, ОПК-1
8.	Применить на практике регрессию	ПК-2, ПК-3, УК-1, ОПК-1
9.	Проанализировать проблемы, возникающие при решении задачи классификации, и предложить пути их решения	ПК-3, УК-1
10.	Проанализировать проблемы, возникающие при решении задачи регрессии, и предложить пути их решения	ПК-3, УК-1
11.	Использовать STL-разложение временного ряда	ПК-1, ПК-2, ОК-2

1.3 Категории обучающихся: уровень образования – ВО, область профессиональной деятельности – для начинающих разработчиков и аналитиков данных.

1.4 Форма обучения – по очному принципу (с применением электронного обучения и дистанционных образовательных технологий) в формате вебинаров.

1.5 Объем программы – 100 академических часов.

Режим обучения – два раза в неделю по 4 академических часа.

2. Содержание программы

2.1. Учебный (тематический) план

Таблица 4

№ п/п	Наименование модулей	Всего, час.	в том числе:				
			лекции	практические занятия (семинары)	самостоятельная работа	контрольные задания	Форма контроля
1	Знакомство с Python (часть 1)	7	2	2	3	0	тестирование
2	Управление вычислениями Контейнеры, итераторы, генераторы.	8	2	2	3	1	тестирование
3	Введение в анализ данных помощью pandas.s	7	2	2	3	0	тестирование
4	Модули для визуализации данных	8	2	2	3	1	тестирование
5	Оптимизация кода с помощью numru, обзор модуля. Работа со строками.	7	2	2	3	0	тестирование
6	Функциональное программирование в Python	8	2	2	3	1	тестирование
7	ООП в Python	7	2	2	3	0	тестирование
8	Командная строка Unix	8	2	2	3	1	тестирование
9	GIT, профилирование и параллелизм	7	2	2	3	0	тестирование
10	Python и web: обзор	8	2	2	3	1	тестирование
11	Теория вероятностей и статистика данных в Python. Понятия о методологии A/B тестирования	7	2	2	3	0	тестирование
12	Пример использования Python для машинного обучения. Модули scikit- learn, Pytorch	8	2	2	3	1	тестирование
13	Итоговая аттестация	10	0	0	10	0	проект
	Итого:	100	24	24	46	6	

2.2. Учебная программа

Таблица 5

№ п/п	Содержание обучения, наименование и тематика практических занятий (вебинаров), самостоятельных работы	Объем , ак.час.
1	Знакомство с Python (часть 1).	7
	Лекция. Введение в Python. работа в Jupyter Notebook и Google Colab, числовые типы данных, условия и циклы.	2
	Практическое занятие..	2
	Самостоятельная работа.	3
	Выполнение контрольных заданий.	0
2	Управление вычислениями. Контейнеры, итераторы, генераторы	8
	Лекция. Set, dict; enumerate, zip, map; list, dict, set comprehensions; itertools, collections.	2
	Практическое занятие..	2
	Самостоятельная работа.	3
	Выполнение контрольных заданий.	1
3	Введение в анализ данных с помощью pandas	7
	Лекция. Введение в анализ данных в Pandas. Фильтрация строк по условию, группировка данных. Работа с несколькими таблицами. Чтение-запись таблиц в файл, формат хранения данных .csv.	2
	Практическое занятие.	2
	Самостоятельная работа.	3
	Выполнение контрольных заданий.	0
4	Модули для визуализации данных	8
	Лекция. Введение в matplotlib. Визуализация данных при помощи графиков и гистограмм. Другие модули для визуализации.	2
	Практическое занятие.	2
	Самостоятельная работа.	3
	Выполнение контрольных заданий.	1
5	Оптимизация кода с помощью numru, обзор модуля. Работа со строками.	7
	Лекция. введение в Numru, оптимизация кода, работа со строками.	2

	Практическое занятие.	2
	Самостоятельная работа.	3
	Выполнение контрольных заданий.	0
6	Функциональное программирование в Python	8
	Лекция.	2
	Практическое занятие.	2
	Самостоятельная работа.	3
	Выполнение контрольных заданий.	1
7.	ООП в Python	7
	Лекция. Объявление класса, создание экземпляра. Атрибуты и методы класса, приватность. Наследование. Множественное наследование. Ромбовидное наследование. Приватность атрибутов. Объектно-ориентированный подход к созданию итераторов и генераторов.	2
	Практическое занятие.	2
	Самостоятельная работа.	3
	Выполнение контрольных заданий.	0
8.	Командная строка Unix	8
	Лекция. – Введение в командную строку. Взаимодействие с оболочкой, навигация и подключение программ. Скрипты; bash. Инструменты оболочки, поиск: файлов; содержимого; команд. Vim, режимы работы. Вставка текста, буферы, вкладки и окна. Быстрые команды и макросы. Работа с процессами. Терминальные мультиплекторы. Alias. SSH, генерация ключа, копирование файлов. Port Forwarding.	2
	Практическое занятие.	2
	Самостоятельная работа.	3
	Выполнение контрольных заданий.	1
9	GIT, профилирование и параллелизм.	7
	Лекция. Версионный контроль git. Модель данных git. Командная строка git. Дебаг и логирование. Профилирование по времени и памяти. Джит компиляция и параллелизм в python. Метапрограммирование: системы сборки, зависимостей. Тестирующие системы (CI).	2
	Практическое занятие.	2
	Самостоятельная работа.	3
	Выполнение контрольных заданий.	0

10	Python и web: обзор.	8
	Лекция. Протокол HTTP, Домены, URL, HTTP-сообщения, основы HTML. Парсинг данных с помощью beautiful soup, прокси. Парсинг и автоматизация взаимодействие с веб с помощью selenium. Быстрое прототипирование и деплоинг мл-моделей со streamlit на примере простейшего классификатора интентов. Обзор примеры работы Flask, верстка с bootstrap.	2
	Практическое занятие.	2
	Самостоятельная работа.	3
	Выполнение контрольных заданий.	1
11	Теория вероятностей и статистика данных в Python. Понятия о методологии A/B тестирования	7
	Лекция.	2
	Практическое занятие.	2
	Самостоятельная работа.	3
	Выполнение контрольных заданий.	0
12	Пример использования Python для машинного обучения. Модули scikit-learn, Pytorch.	8
	Лекция.	2
	Практическое занятие.	2
	Самостоятельная работа.	3
	Выполнение контрольных заданий.	1
13	Итоговая аттестация.	10
	Итого	100

3. Формы аттестации и оценочные материалы

Оценка качества освоения программы проводится по пятибалльной системе по результатам промежуточного контроля (тестирование, проверочные задания на взаимную оценку), контроля посещаемости практических занятий (вебинаров) и результатам итоговой аттестации.

Слушатель считается аттестованным в случае положительных результатов работы (не менее 70% баллов от итоговой оценки) в процессе обучения и успешной сдачи экзамена. При этом баллы за экзамены начисляются только при достижении 50% порога при прохождении каждого экзаменационного испытания. После аттестации слушатель получает оценку “отлично”, если набрано не меньше 80% баллов от возможного максимума, “хорошо”, если набрано не меньше 65%, “удовлетворительно”, если набрано не меньше 40%, в противном случае слушатель курс не сдает.

Результат тестирования, решения проверочных заданий и написания кода проверяется автоматически системой на образовательной платформе. Экзаменационная работа проверяется преподавателем.

Составляющие процесса обучения, которые оцениваются в ходе обучения, и их вклад в итоговую оценку представлены в таблице 6

Таблица 6 – Составляющие процесса обучения

№ п/п	Основные показатели оценки	Вклад в итоговую оценку
1	Основной курс обучения на образовательной платформе	50%
2	Практические занятия	10%
3	Итоговая аттестация	40%

Оценочные материалы

Пример тестового задания

1. Множество из n шаров случайно раскладывают по m ящикам. Найдите вероятность того, что все ящики будут непустыми, если все шары одинаковы.
2. Сколькими способами можно выбрать 3 студентов из 20 и выстроить их в шеренгу?
3. Найти производную u в точке A по направлению вектора a , если

$$u = x^2 + 2xy + y^2 + z^2$$

$A(1;1;1)$

$a(2;-1;0)$

$4/\sqrt{5}$

4. Какое вложение типов верно?
 - a. `integer <float <long`
 - b. `float <integer <long`
 - c. `double <integer <float`
 - d. никакое из перечисленных
5. Что напечатает программа?

```
a=10
```

```
b =3
```

```
while (a > b):
```

```
    a = a//b
```

```
print(b//a)
```

здесь // -- целочисленное деление

6. Пусть дан отсортированный массив. Какая сортировка будет работать быстрее?
 - a. Быстрая сортировка. На любом массиве она работает за $O(n \log(n))$, а сортировка вставками принадлежит к классу сортировок, работающих за $O(n)$.
 - b. Сортировка вставками. На отсортированном массиве она работает также за $O(n \log(n))$, но требует меньше затрат по памяти.
 - c. Одинаково. На отсортированном массиве быстрая сортировка вырождается в сортировку вставками.
 - d. Сортировка вставками. Алгоритм пройдет по отсортированному массиву по внешнему циклу и не будет заходить внутрь условного
7. В каком случае можно использовать бинарный поиск?
 - a. Если мы работаем с массивом, состоящим только из нулей и единиц
 - b. Если массив отсортирован

- c. Если массив можно перевести в бинарную запись
- d. Если к массиву применим бинарный сдвиг
- 8. Можно ли отсортировать массив за линейное время?
 - a. Да, всегда
 - b. Да, если мы можем разбить каждый элемент на разряды
 - c. Нет, никогда
 - d. Нет, если мы можем разбить каждый элемент на разряды

4. Организационно-педагогические условия реализации программы

4.1 Учебно-методическое обеспечение и информационное обеспечение программы

1. Hastie T., Tibshirani R., Friedman J. The Elements of Statistical Learning. Springer, 2009.
2. Murphy K.P. Machine Learning: A Probabilistic Perspective. The MIT Press, 2012.
3. Sheldon A. Linear Algebra Done Right. Springer, 2015.
4. DasGupta A. Probability for statistics and machine learning. Springer, 2011.
5. Иванов Г.Е. Лекции по математическому анализу. Ч.1 - М.:МФТИ, 2004, 2011.
6. Умнов. Аналитическая геометрия и линейная алгебра (2011) — МФТИ.
7. Нестеров. Методы выпуклой оптимизации (2010)
8. Diez, Barr, Çetinkaya-Rundel, Dorazio. Advanced High School Statistics (2015)
9. DasGupta. Probability for Statistics and Machine Learning: Fundamentals and Advanced Topics (2011)
10. Enders, W., 2003, Applied Econometric Time Series, Wiley Publ., 2nd ed.
11. Vance Martin, Stan Hurn, David Harris. Econometric Modelling with Time Series.
12. Specification, Estimation and Testing. Cambridge University Press, 2013.
13. Прикладная и компьютерная лингвистика / Под ред. И. С. Николаева, О. В. Митрениной, Т. М. Ландо. — М.: URSS, 2016. — 320 с.
14. Автоматическая обработка текстов на естественном языке и компьютерная лингвистика : учеб. пособие / Большакова Е.И., Клышинский Э.С., Ландэ Д.В., Носков А.А., Пескова О.В., Ягунова Е.В. — М.: МИЭМ, 2011. — 272 с.
15. Введение в информационный поиск / Кристофер Д. Маннинг, Прабхакар Рагхаван, Хайнрих Шютце. — М.: Вильямс, 2011. — 528 с.
16. Natural Language Processing for the Working Programmer / Daniël de Kok, 2011.

4.2 Материально-технические условия реализации программы

Таблица 7

Наименование специализированных аудиторий, кабинетов, лабораторий	Вид занятий	Наименование оборудования, программного обеспечения
---	-------------	---

Система дистанционного обучения провайдера массовых открытых онлайн курсов	Лекции	Слушателю необходимо наличие доступа в сеть интернет, компьютер. Преподавателю курса необходимо наличие доступа администратора курса на LMS-платформе к материалам курса.
Информационно-коммуникационная платформа дистанционных семинаров	Практические занятия (дистанционные семинары)	Слушателю необходимо наличие доступа в сеть интернет, компьютер. Преподавателю курса необходимо оборудование для проведения дистанционных семинаров (вебинаров), качественный отказоустойчивый доступ в сеть интернет.
Система дистанционного обучения провайдера массовых открытых онлайн курсов	Самостоятельная работа	Наличие компьютера и доступа в сеть интернет.
Система дистанционного обучения провайдера массовых открытых онлайн курсов	Рубежный контроль, Итоговая аттестация	Наличие компьютера и доступа в сеть интернет.

5. Организация образовательного процесса

Слушатели получают доступ к электронным учебным материалам посредством ресурсов поддержки электронного обучения МФТИ и партнерских образовательных площадок. Форматы представления электронных учебных материалов: в виде массовых онлайн курсов (МООС) в системе дистанционного обучения провайдера массовых открытых онлайн курсов.

Преподаватель проводит практические занятия дистанционно в форме вебинаров с использованием платформы ZOOM (или аналогичной).

Самостоятельная работа выполняется слушателем в удобном для него режиме.

В таблице 8 описаны образовательные технологии.

Таблица 8

№ п/п	Вид занятия	Форма проведения занятий	Цель
1	Лекция	Самостоятельный просмотр видеолекций	Ознакомление слушателей с базовым материалом по тематике курса
2	Практические занятия	Выполнение практических заданий, получение обратной связи от преподавателя. Обсуждение вопросов, возникших в	Практическое освоение теоретических знаний, а также углубление знаний по курсу

		результате просмотра видеолекций и изучения литературы.	
3	Самостоятельная работа	Самостоятельное изучение дополнительных материалов и литературы. Выполнение тренировочных тестов и заданий.	Углубление знаний по курсу.
4	Выполнение контрольных заданий	Выполнение тестов, проверочных заданий. Написание кода на языке Python.	Практическое освоение теоретических знаний, контроль освоения материалов.
5	Итоговая аттестация	Защита выпускной квалификационной работы.	Практическое освоение теоретических знаний, контроль освоения материалов.

6. Составители программы:

Райгородский Андрей Михайлович

Доктор физико-математических наук, директор ФПМИ МФТИ

Благодарный Евгений Владимирович

заведующий учебно-методической лабораторией инноватики ФПМИ

Иванова Анастасия Сергеевна

руководитель проектов учебно-методической лаборатории инноватики ФПМИ

Согласовано

Зам. директора ЦДПО

У.Б. Вещезерова

Согласовано

Директор ФПМИ, д.ф.-м.н.

А.М. Райгородский