

МИНИСТЕРСТВО ПРОСВЕЩЕНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
Министерство образования Московской области

Автономная некоммерческая общеобразовательная организация
«Физтех-лицей» имени П. Л. Капицы

Принята на заседании
педагогического совета
Протокол № _____
от «_____» _____ 202__ г.

Утверждена
Приказом № _____ от
«_____» _____ 202__ г.
Директор
Машкова М.Г.

**Дополнительная общеобразовательная общеразвивающая программа
«Python и инструменты машинного обучения»**

Направленность программы: техническая

Уровень программы: средний

Возраст обучающихся: 16-18 лет

Срок реализации программы: 1 год

Руководитель направления:

Райгородский Андрей Михайлович
Доктор физико-математических наук,
директор ФПМИ МФТИ

Авторы-составители:

Благодарный Евгений Владимирович
заведующий учебно-методической
лабораторией инноватики
Иванова Анастасия Сергеевна
руководитель проектов учебно-методической
лаборатории _____ инноватики
Склонен Илья, МФТИ, аспирант кафедры
информатики и вычислительной математики

г.
2023

Долгопрудный

Обратная сторона титульного листа:

Программа прошла внутреннюю экспертизу и рекомендована к реализации в АНОО «Физтех-лицей» имени П. Л. Капицы

Экспертное заключение (рецензия) № _____ от « ____ » _____ 202__ г.

Эксперт (ФИО, должность)

Раздел 1. Пояснительная записка

1.1. Нормативно-правовые основы разработки ДООП.

1.2. Направленность ДООП научная.

Дополнительная общеразвивающая программа «Python и инструменты машинного обучения» — технической направленности.

1.3. Актуальность программы.

Согласно Указу о национальных целях развития России до 2030 года (от 21 июля 2020 года) одними из задач являются возможности для самореализации и развития талантов, цифровая трансформация. И Указу Президента РФ от 1 декабря 2016 г. N 642 «О Стратегии научно-технологического развития Российской Федерации» одним из перспективных направлений является переход к передовым цифровым, интеллектуальным производственным технологиям, роботизированным системам, новым материалам и способам конструирования, создание систем обработки больших объемов данных, машинного обучения и искусственного интеллекта.

Отличительные особенности программы — формирование эффективной системы выявления, поддержки и развития способностей и талантов у детей и молодежи, основанной на принципах справедливости, всеобщности и направленной на самоопределение и профессиональную ориентацию всех обучающихся.

Уровень освоения программы — средний в соответствии с Методическими рекомендациями по проектированию дополнительных общеразвивающих программ (включая разноуровневые программы) министерства образования и науки РФ (письмо от 18 ноября 2015 г. № 09-3242).

1.5. Цель и задачи программы.

Цель программы — первичная подготовка начинающего Data Scientist для которого необходимо знание языка программирования Python.

Целевые ориентиры дополнительного образования детей определены в документах:

- Федеральный закон «Об образовании в Российской Федерации» от 29.12.2012 N 273-ФЗ (ст.2 п.1, 2, 3, 14; ст. 75);
- Приказ Министерства просвещения РФ от 27 июля 2022 г. N 629 «Об утверждении Порядка организации и осуществления образовательной деятельности по дополнительным общеобразовательным программам»;
- Концепция развития дополнительного образования детей (утверждена распоряжением Правительства РФ от 31 марта 2022 года N 678-р).

Задачи	обучения:
Участники будут владеть навыками:	
- обработки данных из разных источников, анализа показателей и принятия стратегических решений на основе данных;	
- анализа больших данных с использованием существующей в организации методологической и технологической инфраструктуры;	
- выявления истинных бизнес-проблем.	

Задачи	воспитания:
- подготовить к самостоятельному выбору профессиональных интересов с помощью реализации разных проектных задач;	
- создать условия для социализации и самореализации обучающихся.	

Задачи	развития:
- развивать мотивационные качества учащихся, мотивация к деятельности и решению проблем-задач;	
- развивать интеллектуальные качества учащихся, познавательный интерес и способности, используя данные о применении изучаемых явлений в окружающей жизни, о новостях науки и техники, учебные дискуссии.	

1.6. Категория обучающихся

Обучающийся:

- 16-18 лет;
- владеет навыками программирования на любом языке, знание математики за пределами школьной программы;
- проходит два этапа отбора: тестирование и защита работ. Допускается дополнительный набор обучающихся на следующий год обучения на основании результатов тестирования.

1.7. Сроки реализации и объем программы.

Срок реализации программы – 1 год. Объем программы — 144 часа.

1.8. Формы организации образовательной деятельности и режим занятий.

Формы организации образовательной деятельности: индивидуальные, групповые. Виды занятий: лекции, практические и семинарские занятия, лабораторные работы, круглые столы, мастер-классы, мастерские, деловые и ролевые игры, тренинги, выездные тематические занятия, выполнение самостоятельной работы, концерты, выставки, творческие отчеты, соревнования и другие виды учебных занятий и учебных работ.

Указывается продолжительность и режим занятий в неделю, количество учебных часов в год. Режим занятий должен соответствовать заявленному уровню программы и требованиям Санитарно-эпидемиологических правил и нормативов СанПиН 2.4.4.3172-14.

Численный состав объединения определяется локальным актом образовательной организации в соответствии с приказом Министерства просвещения РФ от 27 июля 2022 г. N 629 «Об утверждении Порядка организации и осуществления образовательной деятельности по дополнительным общеобразовательным программам».

Пример. Занятия проводятся в разновозрастных группах, численный состав группы – 15 человек.

Формы организации образовательной деятельности – групповые, индивидуальные.

Виды занятий: теоретические и практические занятия, деловые и ролевые игры, выставки, творческие отчеты.

Режим занятий: 1-й год обучения — занятия проводятся 2 раза в неделю по 2 часа, 2-й год обучения – 3 раза в неделю по 2 часа.

Продолжительность занятий – 45 минут, перерыв 10 минут.

1.9. Планируемые результаты освоения программы.

Метапредметные результаты освоения программы:

Обучающиеся будут:

- с помощью языка программирования выполнять сложные проекты, приближенные к реальным задачам в индустрии;
- уметь оценивать результат своей деятельности и достигать результата и извлекать пользу даже из отрицательного результата.

Личностные результаты освоения программы:

	п/п	Обучающийся должен обладать следующим:
1	Научно-исследовательская деятельность	способностью понимать, совершенствовать и применять современный математический аппарат.
2	Проектная и производственно-технологическая деятельность	способностью к разработке и применению алгоритмических и программных решений в области системного и прикладного программного обеспечения.

Результаты воспитания:

У обучающихся будут сформированы:

- устойчивый познавательный интерес к профессии Data Scientist;
- ориентация на достижение успеха и перевод проблемы в задачу;
- готовность к самосовершенствованию, духовному саморазвитию.

Раздел 2. Содержание программы.

Содержание программы соотносится с целью и планируемыми результатами ее освоения.

Нормативные ориентиры к отбору содержания дополнительных общеразвивающих программ определены в документах:

- Приказ Министерства просвещения РФ от 27 июля 2022 г. N 629 «Об утверждении Порядка организации и осуществления образовательной деятельности по дополнительным общеобразовательным программам»;
- Концепция развития дополнительного образования детей (утверждена распоряжением Правительства РФ от 31 марта 2022 года N 678-р).

№ п/п	Наименование модулей	Всего, час.	в том числе:				
			лекции	практические занятия (семинары)	самостоятельная работа	контрольные задания	Форма контроля
1	Знакомство с Python	11	2	5	3	1	тестирование
2	Управление вычислениями. Контейнеры, итераторы, генераторы.	11	2	5	3	1	тестирование
3	Введение в анализ данных с помощью pandas	11	2	5	3	1	тестирование
4	Pandas	11	2	5	3	1	тестирование
5	Модуль Numpy	11	2	5	3	1	тестирование
6	Модули для визуализации данных	11	2	5	3	1	сдача этапа одного из этапов проекта
7	Функциональное программирование в Python	11	2	5	3	1	тестирование
8	Классы в Python	11	2	5	3	1	тестирование
9	Объектно-ориентированный подход к созданию итераторов и генераторов	11	2	5	3	1	сдача этапа одного из этапов проекта
10	Исключения, модули, встроенный модуль	11	2	5	3	1	сдача этапа одного из этапов проекта
11	Python и web: обзор	11	2	5	3	1	тестирование
12	Python для машинного обучения	11	2	6	3	1	тестирование

13	Итоговая аттестация	11	0	0	10	1	проект
	Итого:	144	24	61	46	13	

2.1. Учебный (тематический) план на 1 год обучения:

2.2. Содержание учебного плана:

№ п/п	Содержание обучения, наименование и тематика практических занятий (вебинаров), самостоятельных работы	Объем , ак.час.
1	Знакомство с Python (часть 1).	11
	Лекция. Введение в Python. работа в Jupyter Notebook и Google Colab, числовые типы данных, условия и циклы.	2
	Практическое занятие.	5
	Самостоятельная работа.	3
	Выполнение контрольных заданий.	1
2	Управление вычислениями. Контейнеры, итераторы, генераторы	11
	Лекция. Set, dict; enumerate, zip, map; list, dict, set comprehensions; itertools, collections.	2
	Практическое занятие..	5
	Самостоятельная работа.	3
	Выполнение контрольных заданий.	1
3	Введение в анализ данных с помощью pandas	
	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.	5
	Самостоятельная работа.	3

	Выполнение контрольных заданий.	1
4	Pandas	
	Лекция. API pandas+matplotlib для более компактного построения графиков - Первичный анализ и очистка данных - Фильтрация строк по условию, группировка данных	2
	Практическое занятие.	5
	Самостоятельная работа.	3
	Выполнение контрольных заданий.	1
5	Модуль Numpy	
	Лекция. введение в Numpy, оптимизация кода, работа со строками.	5
	Практическое занятие.	3
	Самостоятельная работа.	1
	Выполнение контрольных заданий.	2
6	Модули для визуализации данных	
	Лекция. Введение в matplotlib. Визуализация данных при помощи графиков и гистограмм. Другие модули для визуализации	2
	Практическое занятие.	5
	Самостоятельная работа.	3
	Выполнение контрольных заданий.	1
7.	Функциональное программирование в Python	
	Лекция. Базовый синтаксис: объявление/вызов, аргументы по умолчанию. Рекурсия. Передача по ссылке. Конструкции args, kwargs. Распаковка контейнеров при вызове. Генераторы. Области видимости. Замыкания. Объектная природа функций. Декораторы. Лямбда-функции	2
	Практическое занятие.	5
	Самостоятельная работа.	3

	Выполнение контрольных заданий.	1
8.	Классы в Python	
	Лекция. Основные принципы ООП. Объявление класса, создание экземпляра. Атрибуты и методы класса. Наследование. Множественное наследование. Ромбовидное наследование. Приватность атрибутов	2
	Практическое занятие.	5
	Самостоятельная работа.	3
	Выполнение контрольных заданий.	1
9	Объектно-ориентированный подход к созданию итераторов и генераторов	
	Лекция. Объектно-ориентированный подход к созданию итераторов и генераторов. Магические методы. Перегрузка операторов	2
	Практическое занятие.	5
	Самостоятельная работа.	3
	Выполнение контрольных заданий.	1
10	Исключения, модули, встроенный модуль	
	Лекция. Устройство исключений, создание и работа с модулями, встроенный модуль python, работа с файловой системой	2
	Практическое занятие.	5
	Самостоятельная работа.	3
	Выполнение контрольных заданий.	1
11	Python и web: обзор.	
	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 интенгов. Обзор примеры работы Flask, верстка с bootstrap.	2

	Практическое занятие.	5
	Самостоятельная работа.	3
	Выполнение контрольных заданий.	1
12	Python для машинного обучения.	
	Лекция. Модули scikit-learn, Pytorch.	2
	Практическое занятие.	6
	Самостоятельная работа.	3
	Выполнение контрольных заданий.	1
13	Итоговая аттестация.	11
	Итого	144

Раздел 3. Методическое обеспечение ДООП.

3.1. 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 и результатам итогового проекта.

Обучающийся считается аттестованным в случае положительных результатов работы (не менее 70% баллов от итоговой оценки) в процессе обучения и успешной защите проекта. При этом баллы за защиту проекта начисляются только при достижении 50% порога при прохождении каждого этапа реализации проекта. После обучения обучающий 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 программа не освоена.

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

Составляющие процесса обучения, которые оцениваются в ходе обучения, и их вклад в итоговую оценку представлены в таблице 1.

Таблица 1 – Составляющие процесса обучения

№ п/п	Основные показатели оценки	Вклад в итоговую оценку
1	Посещение занятий в технопарк	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. $\text{integer} \subset \text{float} \subset \text{long}$

b. $\text{float} \subset \text{integer} \subset \text{long}$

c. $\text{double} \subset \text{integer} \subset \text{float}$

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

3.2. Материально-техническое обеспечение программы.

Оборудованные учебные кабинеты, лаборатории, актовые залы и коворкинги, конференц-залы. Технические средства, используемые в образовательном процессе: компьютерное и мультимедийное оборудование, видео-, аудиовизуальные средства, специальное профессиональное оборудование для реализации учебных проектов.

№ п/ п	Вид занятия	Форма проведения занятий	Цель
1	Лекция	Самостоятельный просмотр видеолекций	Ознакомление с базовым материалом по тематике.

2	Практические занятия	Выполнение практических заданий, получение обратной связи от преподавателя. Обсуждение вопросов, возникших в результате просмотра видеолекций и изучения литературы.	Практическое освоение теоретических знаний, а также углубление знаний.
3	Самостоятельная работа	Самостоятельное изучение дополнительных материалов и литературы. Выполнение тренировочных тестов и заданий.	Углубление знаний.
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.

3.3. Кадровое обеспечение программы.

Программа может быть реализована педагогом дополнительного образования с уровнем образования и квалификации, соответствующим Профессиональному стандарту «Педагог дополнительного образования детей и взрослых».

3.4. Список литературы:

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.