

**Автономная Некоммерческая Организация
Дополнительного Профессионального Образования
«Образовательные технологии Яндекса»**

Утверждено _____

От «_____» _____ 2026 года

Дополнительная общеобразовательная общеразвивающая программа

«Введение в искусственный интеллект»

Объем (трудоемкость) программы: 48 часов

Москва, 2026 год

**Автономная Некоммерческая Организация
Дополнительного Профессионального Образования
«Образовательные технологии Яндекса»**

Утверждено _____

От « _____ » _____ 2026 года

**Дополнительная общеобразовательная общеразвивающая программа
«Введение в искусственный интеллект»**

Вид дополнительной общеобразовательной программы: дополнительная общеразвивающая программа

Объём (трудоемкость) программы: 48 часов

Форма реализации ДООП (форма организации учебного процесса): очная

Категория слушателей: учащиеся школ и студенты средних специальных учебных заведений в возрасте от 13 до 20 лет

Форма обучения: очная-заочная с применением дистанционных технологий

Контроль за выполнением программы возлагается на руководителя программы

Режим занятий: 4 академических часа в неделю

Язык ДООП: обучение производится на русском языке

Руководитель программы _____ / Паволоцкий А.В.

**Автономная Некоммерческая Организация
Дополнительного Профессионального Образования
«Образовательные технологии Яндекса»**

Утверждено _____

От «_____» _____ 2026 года

**Дополнительная общеобразовательная общеразвивающая программа
«Введение в искусственный интеллект»**

Объём (трудоемкость) программы: 48 часов

УЧЕБНЫЙ ПЛАН

№	Наименование разделов	Всего часов	Содержание	Форма контроля
1	Модуль 1. Введение в искусственный интеллект	4	Что такое AI, ML и большие языковые модели. Современные нейронные сети (Алиса, DeepSeek). Промпт-инжиниринг: роль, задача, данные, требования. Контекст и системный промпт. Знакомство со средой LMS Яндекс Лицея.	Создание промптов. Решение задач по пройденным темам.
2	Модуль 2. Электронные таблицы и анализ данных	6	Структура табличных данных: объекты, признаки. Формулы и функции в электронных	Практическая работа: анализ датасета «Кинопоиск-250» средствами

			таблицах. Фильтрация, сортировка, сводные таблицы. Базовая визуализация. Датасет «Кинопоиск-250».	электронных таблиц.
3	Модуль 3. Минимальный Python и анализ данных с pandas	14	Среда Jupyter. Переменные, типы данных, условия, циклы, списки, функции. Чтение CSV в pandas, методы .head(), .info(), .describe(). Фильтрация, сортировка, группировка, агрегация. Базовая визуализация. Формулирование наблюдений по данным.	Практическая работа: разведывательный анализ данных датасета «Кинопоиск-250» средствами pandas.
4	Модуль 4. Введение в машинное обучение. Классификация	12	Что такое машинное обучение: правила, выученные машиной. Типы задач, признаки и целевая переменная. Понятие baseline и train/test split. Метрика accuracy и её ограничения.	Практическая работа: «Классификация на датасете Titanic».

			Постановка задачи классификации. Логистическая регрессия. Метрики Precision, Recall, F1-score. Датасет Titanic.	
5	Модуль 5. Обработка естественного языка (NLP)	6	Текст как данные. Очистка и токенизация. Подход Bag-of-Words и TF-IDF (без формул). Идея распределённых представлений слов (эмбединги). Обзор современных языковых моделей. Применение готовых моделей Hugging Face: классификация тональности, zero-shot, NER.	Практическая работа: анализ текстовых данных с применением готовых моделей.
	Резервное занятие	6	Повторение и разбор сложных тем.	
	ВСЕГО:	48		

Руководитель программы _____ / Паволоцкий А.В.

**Автономная Некоммерческая Организация
Дополнительного Профессионального Образования
«Образовательные технологии Яндекса»**

Утверждено _____

От «_____» _____ 2026 года

РАБОЧАЯ ПРОГРАММА

дополнительной общеобразовательной общеразвивающей программы
«Введение в искусственный интеллект»

Разработчики программы:

Хархота А.Д.

Паволоцкий А.В.

1. ЦЕЛЬ, НОВИЗНА И АКТУАЛЬНОСТЬ

1.1. Цель ДООП:

Создание условий для первичного знакомства с искусственным интеллектом, машинным обучением и большими языковыми моделями; формирование практических навыков работы с табличными и текстовыми данными средствами электронных таблиц, языка программирования Python и готовых сервисов на базе ИИ; подготовка к осознанному применению технологий ИИ в учебной и проектной деятельности.

1.2. Задачи ДООП:

- Формирование понимания основных концепций искусственного интеллекта, машинного обучения и больших языковых моделей.
- Развитие навыков осознанного взаимодействия с LLM-сервисами (Алиса, DeepSeek): составление промптов, учёт контекста, критическая оценка ответов.
- Освоение базовых инструментов анализа табличных данных в электронных таблицах.
- Освоение минимального набора конструкций языка Python и библиотеки pandas, достаточного для исследования реальных табличных датасетов.
- Знакомство с задачей классификации на примере датасета Titanic: подготовка данных, обучение модели, интерпретация метрик качества.
- Знакомство с современными направлениями обработки естественного языка и применением готовых моделей Hugging Face.
- Развитие алгоритмического и аналитического мышления.
- Воспитание критического подхода к интерпретации данных и результатов работы алгоритмов и языковых моделей.

1.3. Актуальность программы:

Современный мир переживает революцию в области искусственного интеллекта. Технологии ИИ и большие языковые модели проникают во все

сферы жизни – от образования и медицины до бизнеса и искусства. При этом школьное образование не успевает за стремительным развитием этой области, что создаёт разрыв между академическими знаниями и реальными потребностями цифровой экономики.

Данная программа призвана сократить этот разрыв, предоставляя учащимся:

- Понимание того, что такое ИИ, машинное обучение и LLM, и каковы их возможности и ограничения.
- Практические навыки осознанного взаимодействия с современными ИИ-сервисами.
- Базовые навыки работы с табличными данными средствами электронных таблиц и Python.
- Первое знакомство с задачей классификации и метриками качества моделей машинного обучения.
- Опыт применения готовых моделей обработки естественного языка к практическим задачам.

Программа разработана с учётом:

- Требований Указа Президента РФ № 490 от 10.10.2019 «О развитии искусственного интеллекта».
- Опыта ведущих технологических компаний в области ИИ-образования.
- Лучших практик преподавания основ машинного обучения для начинающих.

Особенности программы:

- Практико-ориентированный подход: большая часть времени отводится практической работе.
- Использование реальных датасетов (Кинопоиск-250, Titanic) на всём протяжении курса.
- Сквозная работа с большими языковыми моделями: от первого знакомства в Модуле 1 до применения готовых NLP-моделей в Модуле 5.
- Акцент на понимание базовых принципов работы алгоритмов и моделей, а не только их применение.

1.4. Планируемые результаты обучения:

По окончании курса слушатель будет:

Знать:

- Основные концепции искусственного интеллекта, машинного обучения и больших языковых моделей.
- Принципы составления качественных промптов и работы с контекстом в LLM.
- Базовые операции анализа табличных данных в электронных таблицах и в pandas.
- Принцип работы базовой модели классификации (логистическая регрессия) на обзорном уровне.
- Основные метрики качества классификации: Accuracy, Precision, Recall, F1-score.
- Современные подходы к представлению текста: Bag-of-Words, TF-IDF, эмбединги, трансформеры – на обзорном уровне.

Уметь:

- Формулировать промпты для LLM с учётом роли, задачи, данных и требований.
- Проводить базовый разведывательный анализ табличных данных в электронных таблицах и средствами pandas.
- Применять готовый алгоритм классификации к реальному датасету и интерпретировать результат.
- Сравнить качество модели с базовым решением (baseline).
- Применять готовые модели обработки естественного языка к практическим задачам.

Владеть:

- Базовыми навыками работы с электронными таблицами, Jupyter-средой и базовыми библиотеками анализа данных (pandas, scikit-learn).
- Подходами к интерпретации результатов работы моделей машинного обучения.
- Навыками критической оценки ответов больших языковых моделей.

1.5. Уровень ДООП:

Базовый уровень. Программа предполагает первое знакомство с искусственным интеллектом через практическую работу с реальными данными и не требует предварительных знаний программирования или статистики.

2. СОДЕРЖАНИЕ ПРОГРАММЫ

Модуль 1. Введение в искусственный интеллект (4 часа)

Модуль направлен на формирование первоначального представления об искусственном интеллекте, машинном обучении и больших языковых моделях, а также на освоение базовых навыков осознанного взаимодействия с современными LLM-сервисами.

В рамках модуля рассматриваются понятия AI, ML, нейронных сетей и LLM, их применение в реальной жизни и ограничения (галлюцинации, цена ошибки). Обучающиеся знакомятся с современными сервисами Алиса, YandexGPT и DeepSeek, осваивают четырёхчастную модель промпта (роль / задача / данные / требования), приём «режим интервью», понятие контекста и system prompt. Отдельное внимание уделяется экосистеме образовательных продуктов Яндекса (Лицей, ШАД) как пути дальнейшего развития. Практическая часть модуля включает индивидуальное исследование одной из профессий в области ИИ через сравнение ответов двух LLM-сервисов, а также знакомство со средой LMS Яндекс Лицея и решение первых учебных задач.

Темы:

2.1.1. Тема 1: Что такое AI, ML, нейронные сети и LLM. Применения и ограничения.

2.1.2. Тема 2: Промпт-инжиниринг. Контекст и system prompt. Экосистема ИИ-продуктов Яндекса.

2.1.3. Тема 3: Знакомство со средой LMS Яндекс Лицея. Решение первых задач.

Форма проведения промежуточной аттестации: очно-заочная.

Описание, требования к выполнению: обучающиеся выполняют практическую работу, в которой необходимо:

- выбрать одну из профессий в области ИИ;
- составить промпт по четырёхчастной модели и получить ответ от двух LLM-сервисов;
- сравнить ответы и сформулировать выводы;
- оформить результат в виде короткой записи в LMS.

Результат: проанализированные ответы LLM, решённые задачи с автоматической проверкой.

Модуль 2. Электронные таблицы и анализ данных (6 часа)

Модуль посвящён первичному знакомству с табличными данными и их анализу средствами электронных таблиц. Использование электронных таблиц до перехода на Python позволяет освоить базовые операции с данными (фильтрация, сортировка, агрегация, визуализация) на привычном инструменте, прежде чем погружаться в программирование.

В рамках модуля рассматриваются структура датасета (объекты, признаки), формулы и встроенные функции электронных таблиц, фильтрация и сортировка данных, сводные таблицы как способ группировки и агрегации, построение базовых графиков для описания распределений и сравнения категорий. Основным учебным датасетом является «Кинопоиск-250» – рейтинги и характеристики фильмов.

Темы:

2.2.1. Тема 1: Структура табличных данных. Объекты, признаки, типы данных.

2.2.2. Тема 2: Формулы и встроенные функции электронных таблиц.

2.2.3. Тема 3: Фильтрация, сортировка, сводные таблицы.

2.2.4. Практическая работа: исследование датасета «Кинопоиск-250».

Форма проведения промежуточной аттестации: очно-заочная.

Описание, требования к выполнению:

Обучающиеся выполняют практическую работу по анализу датасета «Кинопоиск-250», в которой необходимо:

- описать структуру датасета и типы признаков;
- выполнить набор операций с данными (фильтрация, сортировка, сводные таблицы);
- построить не менее двух графиков;
- сформулировать не менее двух содержательных наблюдений по данным.

Результат: оформленный файл электронной таблицы с результатами анализа.

Модуль 3. Минимальный Python и анализ данных с pandas (14 часов)

Модуль направлен на освоение минимального набора конструкций языка Python и библиотеки pandas, достаточного для самостоятельного исследования реальных табличных датасетов. Главная методическая установка модуля – формирование навыка чтения и понимания кода (а не его написания с нуля), что соответствует реальной практике работы с данными.

Первая часть модуля (3 урока) посвящена базовому Python: переменные, типы данных, условные конструкции, циклы, списки, словари, функции, объектно-методный синтаксис вызовов. Вторая часть (4 урока) посвящена pandas: чтение CSV, первичный осмотр данных (`.head()`, `.info()`, `.describe()`), фильтрация, сортировка, группировка и агрегация, базовая визуализация. Учебной средой модуля является JupyterLab; учебным датасетом – «Кинопоиск-250», что обеспечивает преемственность с Модулем 2.

Темы:

2.3.1. Тема 1: Среда Jupyter. Переменные, типы данных, условные конструкции.

2.3.2. Тема 2: Структуры данных (списки, словари) и циклы.

2.3.3. Тема 3: Функции. Объектно-методный синтаксис как мост к pandas и scikit-learn.

2.3.4. Тема 4: Чтение данных в pandas. Первичный просмотр: `.head()`, `.info()`, `.describe()`.

2.3.5. Тема 5: Фильтрация, сортировка, выбор столбцов.

2.3.6. Тема 6: Группировка и агрегация данных.

2.3.7. Практическая работа: разведывательный анализ датасета «Кинопоиск-250» средствами pandas.

Форма проведения промежуточной аттестации: очно-заочная.

Описание, требования к выполнению:

Обучающиеся выполняют практическую работу средствами pandas, в которой необходимо:

- загрузить датасет «Кинопоиск-250» и описать его структуру;
- выполнить основные операции анализа данных (фильтрация, группировка, агрегация);
- построить не менее двух графиков;
- сформулировать не менее двух содержательных наблюдений по данным;
- оформить результаты в виде структурированного Jupyter-ноутбука.

Результат: рабочий Jupyter-ноутбук с результатами анализа.

Модуль 4. Введение в машинное обучение. Классификация (12 часов)

Модуль посвящён задаче классификации как первой задаче машинного обучения, с которой знакомится обучающийся. Первая часть модуля (2 урока) представляет собой плавное введение в машинное обучение: что такое ML и чем оно отличается от обычной программы, виды задач, понятия признаков и целевой переменной, базовое решение (baseline), концепция разделения данных на обучающую и тестовую выборки. Вторая часть модуля (4 урока) посвящена непосредственно классификации: постановка задачи, обучение модели логистической регрессии, основные метрики качества и их интерпретация.

В качестве основного учебного примера используется датасет Titanic, позволяющий отработать полный цикл решения прикладной задачи классификации: от анализа признаков пассажиров и формулирования простых правил вручную до обучения модели и оценки её качества по метрикам Accuracy, Precision, Recall, F1-score. Отдельное внимание уделяется ограничениям метрики accuracy на несбалансированных данных.

Темы:

2.4.1. Тема 1: Что такое машинное обучение. Виды задач. Признаки и целевая переменная.

2.4.2. Тема 2: Baseline, train/test split, метрика accuracy и её ограничения.

2.4.3. Тема 3: Постановка задачи классификации. Подготовка данных на примере Titanic.

2.4.4. Тема 4: Логистическая регрессия. Обучение базовой модели.

2.4.5. Тема 5: Метрики качества: Precision, Recall, F1-score. Confusion matrix.

2.4.6. Практическая работа: классификация на датасете Titanic.

Форма проведения промежуточной аттестации: очно-заочная.

Описание, требования к выполнению:

Обучающиеся выполняют мини-проект, в котором необходимо:

- подготовить данные датасета Titanic для обучения модели;
- сравнить результат модели с базовым решением (baseline);
- обучить модель логистической регрессии;
- рассчитать основные метрики качества (Accuracy, Precision, Recall, F1-score);
- интерпретировать результат и описать, какие признаки оказались полезными.

Результат: рабочий Jupyter-ноутбук с решением задачи классификации на датасете Titanic.

Модуль 5. Обработка естественного языка (6 часов)

Модуль посвящён основам обработки естественного языка и способам представления текста в виде числовых признаков. Обучающиеся знакомятся с базовыми подходами обработки текстов и получают представление о развитии методов – от простых счётчиков слов до современных контекстных моделей.

В рамках модуля рассматриваются очистка и подготовка текстов, подход Bag-of-Words и TF-IDF (без формул), идея распределённых представлений

слов и эмбедингов на интуитивном уровне, обзор современных языковых моделей и принципа работы трансформеров на базовом уровне. Практическая часть модуля включает применение готовых моделей Hugging Face к реальным задачам: классификация тональности, zero-shot-классификация, извлечение именованных сущностей. Модуль обеспечивает финальный wow-эффект курса за счёт демонстрации возможностей современных предобученных моделей.

Темы:

2.5.1. Тема 1: Текст как данные. Очистка, токенизация. Bag-of-Words и TF-IDF (обзорно).

2.5.2. Тема 2: Распределённые представления слов. Эмбединги. Векторная арифметика на словах.

2.5.3. Тема 3: Современные языковые модели. Применение готовых моделей Hugging Face.

Форма проведения промежуточной аттестации: очно-заочная.

Описание, требования к выполнению:

Обучающиеся выполняют практическую работу по анализу текстовых данных, в которой необходимо:

- подготовить текстовые данные к анализу;
- применить готовую модель Hugging Face к практической задаче (тональность, zero-shot или NER);
- интерпретировать полученный результат;
- продемонстрировать понимание различий между Bag-of-Words, эмбедингами и современными языковыми моделями.

Результат: рабочий Jupyter-ноутбук с решением текстовой задачи на основе готовой модели.

Резервное занятие (6 часов)

Резервное занятие используется для повторения и разбора наиболее сложных тем по запросу группы, дополнительной практики или восполнения пропусков. Размещение резерва внутри расписания определяется преподавателем по итогам прохождения первых модулей.

3. ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ, ПРИМЕНЯЕМЫЕ В РАМКАХ РЕАЛИЗАЦИИ ДООП И ФОРМЫ ОБРАЗОВАТЕЛЬНОЙ ДЕЯТЕЛЬНОСТИ

3.1. В рамках реализации ДООП используются электронное обучение и дистанционные образовательные технологии.

3.2. Образовательная деятельность слушателей предусматривает следующие виды учебных занятий и учебных работ, включая, но не ограничиваясь: лекции, практические и семинарские занятия, самостоятельная работа в LMS и Jupyter-среде, а также другие виды учебных занятий и учебных работ, определённые учебным планом.

4. ФОРМЫ АТТЕСТАЦИИ (ОЦЕНИВАНИЯ) И ОЦЕНОЧНЫЕ МАТЕРИАЛЫ

4.1. Промежуточная аттестация

Промежуточная аттестация проводится в конце каждого модуля и представляет собой решение практических задач, отражающих содержание изученного материала. Такой формат позволяет последовательно оценивать сформированные навыки: от осознанной работы с LLM и базового анализа табличных данных до решения задачи классификации и применения готовых моделей обработки естественного языка. Непрохождение промежуточной аттестации может являться основанием для отказа в дальнейшем обучении по программе.

Промежуточная аттестация проводится по следующим модулям:

Модуль 1. Введение в искусственный интеллект.

Практическая работа «Профессии в ИИ»: сравнение ответов двух LLM-сервисов на единый промпт, краткая записка о выбранной профессии. Дополнительно – решённые учебные задачи в LMS.

Результат: проанализированные ответы LLM, решённые задачи с автоматической проверкой.

Модуль 2. Электронные таблицы и анализ данных.

Практическая работа по анализу датасета «Кинопоиск-250» средствами электронных таблиц: описание структуры данных, фильтрация, сортировка, сводные таблицы, базовая визуализация, формулирование наблюдений.

Результат: оформленный файл электронной таблицы с результатами анализа.

Модуль 3. Минимальный Python и анализ данных с pandas.

Практическая работа по разведывательному анализу датасета «Кинопоиск-250» средствами pandas: чтение данных, первичный просмотр структуры, фильтрация, сортировка, группировка, базовая визуализация, формулирование наблюдений и гипотез.

Результат: рабочий Jupyter-ноутбук с результатами анализа и решённые задачи с автоматической проверкой.

Модуль 4. Введение в машинное обучение. Классификация.

Мини-проект по задаче классификации: подготовка данных, сравнение с baseline, обучение модели логистической регрессии, расчёт и интерпретация основных метрик качества на примере датасета Titanic.

Результат: рабочий Jupyter-ноутбук с решением задачи классификации.

Модуль 5. Обработка естественного языка (NLP).

Практическая работа по анализу текстовых данных с применением готовой модели Hugging Face к одной из задач (классификация тональности, zero-shot, NER); интерпретация результата.

Результат: рабочий Jupyter-ноутбук с решением задачи.

В процессе работы над Модулями 3, 4 и 5 учащиеся используют Jupyter-ноутбуки для решения задач.

5. ОРГАНИЗАЦИОННО-ПЕДАГОГИЧЕСКИЕ УСЛОВИЯ

5.1. Характеристика педагогических кадров, участвующих в реализации программы.

№ п / п	Наименование дисциплин в соответствии с учебным планом	Должность по штатному расписанию	Фамилия, Имя, Отчество	Какое образовательное учреждение профессионального образования окончил, специальность по диплому	Ученая степень и ученое (почетное) звание	Стаж научно-педагогической работы	Штатный, совместитель, иное
1	Дополнительная программа "Введение в искусственный интеллект»						
	"Введение в искусственный интеллект»	Директор по развитию социальных проектов	Медведев Кирилл Владимирович	МГУ имени М.В. Ломоносова	к.ф.-м.н.	20	Штатный
	"Введение в искусственный интеллект»	Академический руководитель школьных программ / руководитель Лицея	Паволоцкий Александр Владимирович	МИЭМ, Вычислительные машины, комплексы и сети	к.п.н	28	Штатный
	"Введение в искусственный интеллект»	Менеджер образовательных продуктов	Хархота Александр Дмитриевич	СибГУТИ, инженер радиотехники	-	3	Штатный

5.2. Учебно-методическое обеспечение учебного процесса.

5.2.1. Нормативные документы

- Указ Президента РФ № 490 от 10.10.2019 «О развитии искусственного интеллекта в Российской Федерации».

- Федеральный закон «Об образовании в Российской Федерации» № 273-ФЗ.

5.2.2. Литература

- Елена Капаца. «Машинное обучение доступным языком».
- Андрей Бурков. «Машинное обучение без лишних слов».

5.2.3. Дополнительная литература

- М. Лутц. Изучаем Python. – СПб.: Диалектика, 2019.
- Уэс Маккинни. Python и анализ данных. – М.: ДМК Пресс, 2020.

5.2.4. Электронные ресурсы

- Материалы и презентации к урокам в LMS Яндекс Лицея.
- Среда JupyterLab для работы с -ноутбуками.
- Сервисы Алиса (alice.yandex.ru), DeepSeek для практики работы с LLM.
- PromptHub Алисы (alice.yandex.ru/prompthub) для знакомства с библиотекой промптов.
- Каталог моделей Hugging Face (huggingface.co) для практики применения готовых NLP-моделей.
- Документация библиотек pandas, scikit-learn.
- Открытые наборы данных (Kaggle, открытый датасет Кинопоиска).

5.2.5. Материально-техническое обеспечение

Совокупность технологий, в том числе технологий электронного обучения, используемых в образовательной организации, включая, но не ограничиваясь: интерактивная онлайн-среда, позволяющая осуществлять обучение в дистанционном формате со встроенными механизмами автопроверки практических работ; JupyterLab или альтернативная среда для запуска Jupyter-ноутбуков и моделей машинного обучения без локальной установки.

Утверждено _____

От «_____» _____ 2026 года

КАЛЕНДАРНЫЙ УЧЕБНЫЙ ГРАФИК
дополнительной общеобразовательной общеразвивающей программы

«Введение в искусственный интеллект»

Обучение по Программе проходит в очно-заочной форме с применением дистанционных технологий, начиная с **12 октября 2026 года**. Слушатель имеет право начать обучение с начала даты курса, присоединившись к Программе единовременно.

Руководитель программы _____ / Паволоцкий А.В.

