

ОАНО «Школа «ЛЕТОВО»

«СОГЛАСОВАНО»
на заседании педагогического совета
Протокол №1 от 25 августа 2025 г.

«УТВЕРЖДЕНО» Приказом Директора
ОАНО «Школа «ЛЕТОВО»
№ 138-ОД от 26 августа 2025

ДОПОЛНИТЕЛЬНАЯ ОБЩЕОБРАЗОВАТЕЛЬНАЯ
ОБЩЕРАЗВИВАЮЩАЯ ПРОГРАММА

Направленность программы: техническая

Название программы развития в Дипломе Летово:

Наука и познание

Название программы:
Машинное обучение для продолжающих

Возраст обучающихся: 9-11 класс

Срок реализации программы: 1 год

Составитель:

Владимир Михайлович Гуровиц

Подразделение:

Кафедра информационных технологий и дизайна

Москва, 2025

1. ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

Актуальность программы. В современном мире машинное обучение становится одной из ключевых технологий, определяющих развитие науки, экономики и образования. Алгоритмы анализа данных и предсказательного моделирования используются в медицине, инженерии, финансах, социальных науках и во множестве других сфер, обеспечивая качественно новый уровень принятия решений и автоматизации процессов. Владение методами машинного обучения становится важнейшей компетенцией современного специалиста, а понимание принципов работы алгоритмов – необходимым элементом общего научного мировоззрения. Кроме того, в настоящее время увеличивается количество олимпиад для школьников, по результатам которых наиболее успешные участники приглашаются на стажировки, конференции и имеют

Машинное обучение в широком смысле можно рассматривать как совокупность методов преобразования данных с целью извлечения знаний и построения моделей, описывающих скрытые закономерности. Эти методы включают в себя математические и статистические инструменты, способы представления информации, а также подходы к валидации и интерпретации результатов. Подобно тому как технология в инженерии обеспечивает преобразование материалов в изделия, машинное обучение обеспечивает преобразование данных в знания, решения и прогнозы.

Актуальность курса обусловлена необходимостью систематизации знаний учащихся в этой области. В школьной и олимпиадной практике часто встречаются задачи анализа данных, прогнозирования, оптимизации и поиска зависимостей. Однако без понимания основ построения моделей, выбора метрик и способов борьбы с переобучением такие задачи становятся трудными для самостоятельного решения. Курс направлен на формирование у школьников целостного взгляда на машинное обучение как на процесс: от постановки задачи и работы с данными до построения и интерпретации модели.

Особое внимание в программе уделено развитию исследовательских и творческих навыков: учащиеся будут не только изучать алгоритмы, но и пробовать себя в решении олимпиадных задач, требующих поиска нестандартных решений и умения выбирать инструменты под конкретную цель. Это создаёт условия для личностного развития, способствует формированию критического мышления, самостоятельности и умению работать в команде. Таким образом, курс машинного обучения отвечает актуальным образовательным запросам, помогает учащимся интегрироваться в современное научно-техническое пространство и готовит их к будущей профессиональной и исследовательской деятельности.

Аспект новизны. Отличительной особенностью данной программы является её практико-ориентированная направленность, основанная на

включении обучающихся в решение олимпиадных и прикладных задач анализа данных. Курс строится вокруг систематизации разрозненных знаний о машинном обучении: учащиеся учатся не только применять отдельные алгоритмы, но и видеть их место в общем пайплайне решения, понимать взаимосвязь между метриками, моделями, признаками и валидацией. Такой подход позволяет не заучивать инструменты, а осваивать логику построения полноценного решения.

Кроме того, программа отличается междисциплинарностью: при изучении методов машинного обучения используются связи с математикой (линейная алгебра, анализ, теория вероятностей), информатикой (программирование, алгоритмы), экономикой и естественными науками (статистический анализ, обработка временных рядов, интерпретация данных). Эти связи формируют целостное представление о современном научно-техническом подходе к данным и готовят школьников к дальнейшему углубленному изучению как в университетских курсах, так и в исследовательской деятельности.

Курс ориентирован на развитие исследовательских и проектных компетенций: обучающиеся выполняют практические задания, участвуют в мини-соревнованиях и разрабатывают собственные решения, которые защищаются в формате проектов. Это способствует формированию у них критического и системного мышления, умению планировать работу, обосновывать выбор методов и интерпретировать результаты.

Важным аспектом новизны программы является её акцент на рефлексии и самоорганизации: школьники учатся самостоятельно планировать процесс решения задачи, выбирать оптимальные стратегии, анализировать ошибки и корректировать подход. Таким образом, курс не только даёт предметные знания и навыки, но и формирует у обучающихся компетенции исследователя и участника научного сообщества, что напрямую повышает их готовность к участию в олимпиадах и к дальнейшему профессиональному развитию.

Направленность программы – техническая.

Цель программы – развитие у обучающихся системных знаний и практических навыков в области машинного обучения через решение олимпиадных и прикладных задач, формирование умений строить полноценные проекты по машинному обучению и применять современные методы анализа данных.

Задачи, решаемые программой:

- обучение базовым понятиям и формирование практических навыков в области машинного обучения и анализа данных;
- повышение мотивации к изучению математики, информатики и статистики как фундаментальных основ машинного обучения;
- вовлечение школьников в исследовательскую и проектную деятельность, подготовку к олимпиадам и соревнованиям;

- приобщение обучающихся к современным технологиям анализа данных, способным помочь им в реализации собственного интеллектуального и творческого потенциала;
- развитие логического и абстрактного мышления, способности к аналитике и поиску закономерностей;
- формирование творческих навыков в постановке и решении нестандартных задач, развитие познавательной активности;
- развитие коммуникативных умений и навыков командной работы при решении проектов и участии в соревнованиях;
- формирование навыков критического мышления, умения интерпретировать результаты и делать выводы;
- создание условий для повышения самооценки обучающихся, уверенности в своих силах и раскрытия личностного потенциала;
- развитие способностей к самоорганизации, целеустремлённости и самостоятельному планированию учебной и исследовательской деятельности.

Адресат программы: учащиеся 15-18 лет.

Форма обучения: очная.

Особенности организации образовательного процесса: основная форма работы на занятии индивидуальная.

Срок реализации программы: 1 год.

Общее количество часов - 54, количество часов в неделю — два академических часа.

Уровень сложности программы

Данная программа относится к продвинутому уровню сложности и предполагает освоение углублённых знаний, умений и навыков в области машинного обучения, систематизацию ранее полученного опыта и формирование компетенций, необходимых для решения олимпиадных и исследовательских задач.

Планируемые результаты освоения программы **(предметные, личностные и метапредметные)**

Освоение содержания программы обеспечивает достижение обучающимися следующих результатов:

Личностных:

- формирование осознанного отношения к собственным действиям и результатам при выполнении учебных и проектных заданий;
- развитие умения ставить цели, планировать деятельность для их достижения, рассчитывать время и соблюдать сроки выполнения задач;
- поддержка любознательности, инициативности и исследовательского интереса при решении разнообразных аналитических и практических задач;
- развитие внимательности, настойчивости, целеустремлённости и готовности преодолевать трудности;
- формирование самостоятельности суждений, умения отстаивать собственную точку зрения, развитие критического и нестандартного мышления;
- освоение норм сотрудничества и правил взаимодействия в учебных и проектных группах;
- развитие коммуникативной компетентности в общении и совместной деятельности, умения работать в команде и презентовать результаты своей работы.

Метапредметных (в соответствии с рубрикаторм школы):

- Умение взаимодействовать с другими в процессе работы (работа в группе);
- Генерирование идей (Творческое мышление);
- Навыки самоорганизации и эффективного управления временем (Самоорганизация).

Предметных:

В результате освоения программы обучающиеся должны *знать*:

- основные понятия и термины в области машинного обучения и анализа данных;
- различия между лосс-функциями и метриками качества, области их применения;
- принципы валидации и методы борьбы с переобучением;
- основы работы с линейными моделями, деревьями решений и ансамблями;
- принципы работы алгоритмов оптимизации (градиентный спуск, стохастический градиентный спуск);
- подходы к обработке и генерации признаков для различных типов данных (временные ряды, координаты, категории и др.);

- методы снижения размерности и визуализации данных;
- особенности организации полного пайплайна задачи машинного обучения: от предобработки данных до финальной модели;
- базовые инструменты для работы с данными и моделями (язык программирования Python, библиотеки sklearn, Pandas, CatBoost);
- основы ответственного применения машинного обучения, включая интерпретацию и проверку корректности решений.

уметь:

- формулировать задачу машинного обучения, выбирать соответствующие метрики и модели;
- проводить предобработку данных, кодировать категории, создавать новые признаки;
- применять методы валидации, строить разбиения по группам и корректно выделять тестовые данные;
- обучать и настраивать линейные модели, деревья решений, ансамбли;
- использовать методы регуляризации, гиперпараметры и аугментации для борьбы с переобучением;
- работать с большими объёмами данных, использовать специальные библиотеки, мини-группы данных и стохастическую оптимизацию;
- визуализировать данные и результаты моделей;
- планировать и реализовывать исследовательскую или проектную задачу по машинному обучению;
- представлять результаты своей работы в устной и письменной форме.

владеть:

- базовыми понятиями и инструментами современного машинного обучения;
- практическими навыками работы с языком программирования Python с основными библиотеками анализа данных и машинного обучения;
- методами оценки качества моделей и их сравнительного анализа;
- приёмами генерации и отбора признаков;
- навыками командной работы над проектами и участия в мини-соревнованиях.

Формы подведения итогов реализации программы

Подведение итогов реализуется в рамках защиты результатов выполнения практических заданий по каждой теме, а также успешному участию в тренировочных соревнованиях.

Формы демонстрации результатов обучения

Представление результатов образовательной деятельности пройдёт в форме презентации собственного решения предложенной исследовательской задачи и последующих ответов, выступающих на вопросы наставника и других учащихся.

2. Содержание программы и тематический план

Программа предполагает постепенное расширение знаний и их углубление, а также приобретение умений в области проектирования, конструирования и изготовления творческого продукта.

Тематический план

Содержание программы по разделам

<i>Наименование разделов и тем</i>	<i>Содержание учебного материала, лабораторные работы и практические занятия, самостоятельная работа обучающихся</i>		<i>Объем часов</i>
<i>Классическое машинное обучение</i>	1.	Вводное занятие. Формат курса. Обзор существующих олимпиад.	4
	2.	Концепция ML-решения как задачи оптимизации. Разбор алгоритма линейной регрессии. Частные производные. Градиентный спуск.	2
	3.	Метрики и функции потерь. Задача регрессии и задача классификации. Дисбаланс классов.	2
	4.	Проблемы больших данных. Различные оптимизации. Использование специализированных библиотек. Подходы к обработке данных.	2
	5.	Зачетное соревнование по классическим библиотекам машинного обучения (numpy, pandas, matplotlib)	4
	6.	Подходы к генерации признаков. Логические операции. Полиномиальные модели. Линейные и нелинейные ядра.	2
	7.	Валидация моделей. Разделение данных на тренировочную и обучающую выборки. Адверсальная валидация. Стратифицированная валидация.	3
	8.	Пайплайн решения задачи машинного обучения. Эконометрические подходы. Построение и обучение моделей.	3
	9.	Зачетное соревнование по классическому машинному обучению	4
<i>Нейронные сети и глубокое обучение</i>	10.	Построение ансамбля модели линейной регрессии. Концепция перцептрона. Обратное распространение ошибки.	4
	11.	Разнообразие слоев нейронных сетей. Нормировка. Случайное зашумление нейронов. Функции активации.	2

12.	Семинар по размерностям. Теория тензоров. Действия над матрицами.	2
13.	Сверточные нейронные сети. Использование нейронных сетей для обработки изображений. You-Only-Look-Once-подходы.	4
14.	Рекуррентные нейронные сети. Последовательности токенов. Обработка текстов. Двухнаправленные рекуррентные подходы.	4
15.	Модель трансформера. Слои внимания. Энкодер и декодер. Использование архитектуры трансформера для обработки картинок. Визуальный трансформер.	4
16.	Способы генерации текста. Эвристики для улучшения качества. Оптимизация инференса.	2
17.	Проецирование векторного представления текста и картинок в одно измерение. CLIP-подобные модели.	2
18.	Генерация изображений. Зашумление и обратный процесс расшумления. Диффузионные модели.	2
	Итого:	54

Раздел 1. Классическое машинное обучение

Вводное занятие. Знакомство с форматом курса, целями и требованиями. Обзор олимпиад и соревнований по анализу данных. Обсуждение типов задач, которые встречаются в олимпиадной практике.

Концепция ML-решения как задачи оптимизации. Рассмотрение линейной регрессии как базовой модели. Введение понятия функции потерь, частных производных и градиентного спуска. Демонстрация процесса обучения модели и визуализация изменения параметров.

Метрики и функции потерь. Отличие задач регрессии и классификации. Разбор наиболее часто используемых метрик. Обсуждение проблемы дисбаланса классов и способов её учёта при выборе метрики.

Проблемы больших данных. Различие вычислений на CPU и GPU. Использование специализированных библиотек. Методы оптимизации вычислений и подходы к обработке больших объёмов информации. Практика визуализации больших датасетов.

Зачётное соревнование. Работа с классическими библиотеками машинного обучения (NumPy, pandas, matplotlib). Решение прикладной задачи: анализ датасета, построение простых моделей, визуализация и интерпретация результатов.

Генерация признаков. Создание новых признаков через логические операции и полиномиальные преобразования. Разбор линейных и

нелинейных ядер. Проверка влияния дополнительных признаков на качество модели.

Валидация моделей. Принципы разделения данных на тренировочную и тестовую выборки. Методы кросс-валидации, включая стратифицированную. Введение в адверсариальную валидацию и её использование для более корректной оценки качества модели.

Построение пайплайна решения задачи. От исследовательского анализа данных до выбора и обучения моделей. Использование элементов эконометрического подхода. Ознакомление с принципами стекинга моделей.

Финальное зачётное соревнование. Решение комплексной задачи классического машинного обучения. Построение полного пайплайна: исследование данных, обучение и валидация моделей, выбор метрик, интерпретация и защита результатов.

Раздел 2. Нейронные сети и глубокое обучение

Введение в методы объединения моделей для повышения точности предсказаний. Рассмотрение базового перцептрона как простейшей нейронной сети. Объяснение механизма обратного распространения ошибки для обучения многослойных сетей. Демонстрация на простых примерах с визуализацией изменения весов.

Обзор различных типов слоев: полносвязные, свёрточные, рекуррентные. Понятие нормировки и её влияние на стабильность обучения. Использование зашумления нейронов и других методов регуляризации. Разбор функций активации и их эффект на обучение.

Фундаментальные операции с тензорами. Понимание размерностей данных и их влияние на обучение моделей. Примеры матричных операций в контексте нейросетей и оптимизации вычислений. Практические упражнения по трансформации тензоров.

Построение и обучение свёрточных сетей для классификации изображений. Понимание фильтров, свёртки и пулинга. Введение в детекторы объектов на основе YOLO: архитектура, принципы работы и оптимизация для быстрого инференса.

Моделирование последовательных данных с помощью RNN, LSTM и GRU. Применение к задачам обработки текста: прогнозирование следующего токена, анализ последовательностей. Двухнаправленные сети и их преимущества для понимания контекста.

Изучение принципа self-attention и его реализации в энкодерах и декодерах. Применение трансформеров для текста и изображений. Рассмотрение Vision Transformer (ViT) и особенностей работы с визуальными данными.

Методы автогенерации текста: greedy search, beam search, top-k и nucleus sampling. Разбор способов улучшения качества генерации и

предотвращения повторов. Оптимизация инференса для ускорения работы моделей.

Принцип обучения совместного пространства для текста и изображений. Использование таких моделей для поиска и сопоставления мультимодальных данных. Практическая демонстрация на примерах поиска по изображению и тексту.

Введение в диффузионные модели для генерации изображений. Понимание процессов добавления и удаления шума. Примеры обучения и генерации изображений с использованием современных подходов.

3. Формы контроля и оценочные материалы

Промежуточная аттестация и итоговое оценивание будет проводиться с учетом результатов выполнения практических работ.

Формы предъявления и демонстрации образовательных результатов
открытое занятие, проектная деятельность

4. Организационно-педагогические условия реализации программы. **Особенности работы по программе**

В основе образовательного процесса лежит проектно-ориентированный подход. Теоретический материал подается через интерактивные лекции и пошаговые мастер-классы, в которых студенты работают с реальными данными и моделями. Практические задания выполняются индивидуально, в парах или малых группах, включая построение моделей, анализ данных и визуализацию результатов. Занятия проводятся в формате бесед, семинаров и лекций с использованием презентаций, видеодемонстраций и интерактивных инструментов для обучения ML.

Для развития метапредметных навыков применяются инструменты SMART для самоорганизации, работа в малых группах для генерации идей, совместного решения задач и развития навыков коммуникации и взаимодействия в команде.

Содержание обучения	Характеристика основных видов учебной деятельности обучающихся (на уровне учебных действий)
Классическое машинное	В рамках данного раздела обучающиеся: <ul style="list-style-type: none">• Осваивают основные методы машинного обучения, их

обучение	<p>преимущества и недостатки, области применения;</p> <ul style="list-style-type: none"> • Овладевают навыками подготовки данных, построения моделей и визуализации результатов в двух- и многомерных пространствах; • Изучают правила безопасной работы с вычислительными ресурсами и принципами воспроизводимых экспериментов; • Изучают основы работы с библиотеками и инструментами для ML • Разрабатывают алгоритмы и технические пайплайны для решения прикладных задач; • Создают модели для анализа данных и прогноза результатов; • Выполняют самостоятельные и творческие проекты по анализу данных и построению ML-решений.
----------	--

<p>Нейронные сети и глубокое обучение</p>	<ul style="list-style-type: none"> • Знакомятся с разнообразными архитектурами нейронных сетей и их применением в реальных задачах; • Учатся строить и обучать многослойные, сверточные и рекуррентные модели, а также интерпретировать результаты их работы; • Осваивают методы подготовки и нормализации данных, а также визуализации промежуточных слоёв и предсказаний моделей; • Изучают работу современных библиотек и фреймворков для глубокого обучения; • Разрабатывают практические решения для анализа изображений, текста и других типов данных; • Создают собственные проекты, включая классификацию, регрессию и генерацию данных, с последующей оценкой качества моделей; • Выполняют творческие задания, направленные на экспериментирование с архитектурами и настройкой гиперпараметров.
---	---

5. Перечень учебно-методического и материально-технического обеспечения

Материально-техническое оснащение программы

Рабочее место обучающегося:

- личный ноутбук;

Рабочее место наставника:

- Ноутбук – 1 шт.;
- Интерактивный дисплей;
- Маркерная доска;
- Единая сеть Wi-Fi;

Программное обеспечение:

- программное обеспечение для работы с языком Python и интерактивными средами (Python, Jupyter Notebook, JupyterLab);
- программное обеспечение для обработки и анализа данных (NumPy, pandas, OpenCV)
- программное обеспечение для визуализации данных и результатов моделей (matplotlib, seaborn, Plotly);
- программное обеспечение для классического машинного обучения (scikit-learn);
- программное обеспечение для построения и обучения нейронных сетей (PyTorch);
- облачные среды для запуска кода и работы с GPU (Google Colab, Kaggle Notebooks).

Список литературы

1. Браун Л., Чоллетт А. Python для анализа данных. Издательство O'Reilly. 2020 год. 512 с.
2. Комлев А.А. Машинное обучение на Python: от простого к сложному. Москва: БХВ. 2021, 432 с.
3. Хастие Т., Тибширани Р., Фридман Д. Элементы статистического обучения. СПб.: Питер. 2021, 784 с.
4. Гудфеллоу И., Бенджио Й., Курвиль А. Глубокое обучение. СПб.: Питер. 2019, 800 с.
5. Назаров А., Кузнецов П. Введение в нейронные сети и глубокое обучение. Москва: ДМК Пресс. 2020, 360 с.