|  |  |
| --- | --- |
| **Национальный** **исследовательский университет** **«Высшая школа экономики»****Лицей** | **Приложение 656**УТВЕРЖДЕНОпедагогическим советом Лицея НИУ ВШЭпротокол № 13 от 21.06.2024  |
|  |  |

Рабочая программа учебного предмета (курса)
«Основы анализа данных»

10-11 класс

Направление «Информатика, инженерия и математика»

**Автор:**

Палеев Д.А.

1. **Планируемые результаты освоения учебного предмета**

Личностные результаты отражают готовность и способность обучающихся руководствоваться сформированной внутренней позицией личности, системой ценностных ориентаций, позитивных внутренних убеждений, соответствующих традиционным ценностям российского общества, расширение жизненного опыта и опыта деятельности в процессе реализации средствами учебного предмета основных направлений воспитательной деятельности.

1. **гражданского воспитания:**

осознание своих конституционных прав и обязанностей, уважение закона и правопорядка, соблюдение основополагающих норм информационного права и информационной безопасности;

готовность противостоять идеологии экстремизма, национализма, ксенофобии, дискриминации по социальным, религиозным, расовым, национальным признакам в виртуальном пространстве;

1. **патриотического воспитания:**

ценностное отношение к историческому наследию, достижениям России в науке, искусстве, технологиях, понимание значения информатики как науки в жизни современного общества;

1. **духовно-нравственного воспитания:**

сформированность нравственного сознания, этического поведения;

способность оценивать ситуацию и принимать осознанные решения, ориентируясь на морально-нравственные нормы и ценности, в том числе в сети Интернет;

1. **эстетического воспитания:**

эстетическое отношение к миру, включая эстетику научного и технического творчества; способность воспринимать различные виды искусства, в том числе основанного на использовании информационных технологий;

1. **физического воспитания:**

сформированность здорового и безопасного образа жизни, ответственного отношения к своему здоровью, в том числе за счёт соблюдения требований безопасной эксплуатации средств информационных и коммуникационных технологий;

1. **трудового воспитания:**

готовность к активной деятельности технологической и социальной направленности, способность инициировать, планировать и самостоятельно выполнять такую деятельность;

интерес к сферам профессиональной деятельности, связанным с информатикой, программированием и информационными технологиями, основанными на достижениях науки информатики и научно-технического прогресса, умение совершать осознанный выбор будущей профессии и реализовывать собственные жизненные планы;

готовность и способность к образованию и самообразованию на протяжении всей жизни;

1. **экологического воспитания:**

осознание глобального характера экологических проблем и путей их решения, в том числе с учётом возможностей информационно-коммуникационных технологий;

1. **ценности научного познания:**

сформированность мировоззрения, соответствующего современному уровню развития науки, достижениям научно-технического прогресса и общественной практики, за счёт понимания роли информационных ресурсов, информационных процессов и информационных технологий в условиях цифровой трансформации многих сфер жизни современного общества;

осознание ценности научной деятельности, готовность осуществлять проектную и исследовательскую деятельность индивидуально и в группе.

1. **эмоциональный интеллект, предполагающий сформированность:**

 саморегулирования, включающего самоконтроль, умение принимать ответственность за своё поведение, способность адаптироваться к эмоциональным изменениям и проявлять гибкость, быть открытым новому;

внутренней мотивации, включающей стремление к достижению цели и успеху, оптимизм, инициативность, умение действовать, исходя из своих возможностей;

эмпатии, включающей способность понимать эмоциональное состояние других, учитывать его при осуществлении коммуникации, способность к сочувствию и сопереживанию;

социальных навыков, включающих способность выстраивать отношения с другими людьми, заботиться, проявлять интерес и разрешать конфликты.

**Метапредметные результаты:**

В результате изучения информатики на уровне среднего общего образования у обучающегося будут сформированы метапредметные результаты, отраженные в универсальных учебных действиях, а именно – познавательные универсальные учебные действия, коммуникативные универсальные учебные действия, регулятивные универсальные учебные действия, совместная деятельность.

**Познавательные универсальные учебные действия**

*Базовые логические действия:*

самостоятельно формулировать и актуализировать проблему, рассматривать её всесторонне;

устанавливать существенный признак или основания для сравнения, классификации и обобщения;

определять цели деятельности, задавать параметры и критерии их достижения;

выявлять закономерности и противоречия в рассматриваемых явлениях;

разрабатывать план решения проблемы с учётом анализа имеющихся материальных и нематериальных ресурсов;

вносить коррективы в деятельность, оценивать соответствие результатов целям, оценивать риски последствий деятельности;

координировать и выполнять работу в условиях реального, виртуального и комбинированного взаимодействия;

развивать креативное мышление при решении жизненных проблем.

*Базовые исследовательские действия:*

владеть навыками учебно-исследовательской и проектной деятельности, навыками разрешения проблем, способностью и готовностью к самостоятельному поиску методов решения практических задач, применению различных методов познания;

осуществлять различные виды деятельности по получению нового знания, его интерпретации, преобразованию и применению в различных учебных ситуациях, в том числе при создании учебных и социальных проектов; формировать научный тип мышления, владеть научной терминологией, ключевыми понятиями и методами;

ставить и формулировать собственные задачи в образовательной деятельности и жизненных ситуациях;

выявлять причинно-следственные связи и актуализировать задачу, выдвигать гипотезу её решения, находить аргументы для доказательства своих утверждений, задавать параметры и критерии решения;

анализировать полученные в ходе решения задачи результаты, критически оценивать их достоверность, прогнозировать изменение в новых условиях;

давать оценку новым ситуациям, оценивать приобретённый опыт;

осуществлять целенаправленный поиск переноса средств и способов действия в профессиональную среду;

уметь переносить знания в познавательную и практическую области жизнедеятельности;

уметь интегрировать знания из разных предметных областей;

выдвигать новые идеи, предлагать оригинальные подходы и решения, ставить проблемы и задачи, допускающие альтернативные решения.

*Работа с информацией:*

владеть навыками получения информации из источников разных типов, самостоятельно осуществлять поиск, анализ, систематизацию и интерпретацию информации различных видов и форм представления; создавать тексты в различных форматах с учётом назначения информации и целевой аудитории, выбирая оптимальную форму представления и визуализации;

оценивать достоверность, легитимность информации, её соответствие правовым и морально-этическим нормам;

использовать средства информационных и коммуникационных технологий в решении когнитивных, коммуникативных и организационных задач с соблюдением требований эргономики, техники безопасности, гигиены, ресурсосбережения, правовых и этических норм, норм информационной безопасности;

владеть навыками распознавания и защиты информации, информационной безопасности личности.

**Коммуникативные универсальные учебные действия**

*Общение:*

осуществлять коммуникации во всех сферах жизни;

распознавать невербальные средства общения, понимать значение социальных знаков, распознавать предпосылки конфликтных ситуаций и смягчать конфликты;

владеть различными способами общения и взаимодействия, аргументированно вести диалог, уметь смягчать конфликтные ситуации; развёрнуто и логично излагать свою точку зрения с использованием языковых средств.

*Совместная деятельность:*

понимать и использовать преимущества командной и индивидуальной работы; выбирать тематику и методы совместных действий с учётом общих интересов и возможностей каждого члена коллектива;

принимать цели совместной деятельности, организовывать и координировать действия по их достижению: составлять план действий, распределять роли с учётом мнений участников, обсуждать результаты совместной работы;

оценивать качество своего вклада и каждого участника команды в общий результат по разработанным критериям;

предлагать новые проекты, оценивать идеи с позиции новизны, оригинальности, практической значимости;

осуществлять позитивное стратегическое поведение в различных ситуациях, проявлять творчество и воображение, быть инициативным.

**Регулятивные универсальные учебные действия**

*Самоорганизация:*

самостоятельно осуществлять познавательную деятельность, выявлять проблемы, ставить и формулировать собственные задачи в образовательной деятельности и жизненных ситуациях;

самостоятельно составлять план решения проблемы с учётом имеющихся ресурсов, собственных возможностей и предпочтений;

давать оценку новым ситуациям;

расширять рамки учебного предмета на основе личных предпочтений;

делать осознанный выбор, аргументировать его, брать ответственность за решение;

оценивать приобретённый опыт;

способствовать формированию и проявлению широкой эрудиции в разных областях знаний, постоянно повышать свой образовательный и культурный уровень.

*Самоконтроль:*

давать оценку новым ситуациям, вносить коррективы в деятельность, оценивать соответствие результатов целям;

владеть навыками познавательной рефлексии как осознания совершаемых действий и мыслительных процессов, их результатов и оснований, использовать приёмы рефлексии для оценки ситуации, выбора верного решения;

оценивать риски и своевременно принимать решения по их снижению;

принимать мотивы и аргументы других при анализе результатов деятельности.

*Принятия себя и других:*

принимать себя, понимая свои недостатки и достоинства;

принимать мотивы и аргументы других при анализе результатов деятельности;

признавать своё право и право других на ошибку;

развивать способность понимать мир с позиции другого человека.

**Предметные результаты:**

* уметь использовать библиотеки Python для работы с данными: Numpy, Pandas ;
* уметь использовать библиотеки Python для визуализации данных: Seaborn, Matplotlib;
* уметь обрабатывать большие наборы данных;
* знать основные практики “Разведывательного анализа данных” ;
* знать основные понятия и постановки задач машинного обучения;
* уметь анализировать графики и делать выводы по ним;
* уметь воспроизводить принцип проверки статистических гипотез;
* формулировать задачу на языке статистической гипотезы;
* уметь производить проверку статистических гипотез с помощью корректно подобранных статистических критериев и интерпретировать полученные результаты;
* уметь строить линейную регрессионную модель и интерпретировать полученные результаты;
* уметь выбирать оптимальную модель машинного обучения для решения задачи и обосновывать её выбор;
* понимать принцип работы нейронных сетей;
* уметь работать с библиотекой обучения нейронных сетей Pytorch;
* уметь дообучать свёрточные нейронные сети под конкретные задачи;

Учебный курс «Основы анализа данных» ориентирован на учащихся, которые хотят развить практические навыки в сфере анализа данных и машинного обучения. Данная программа рассчитана на углубленное изучение учебного курса в течение 68 часов в 10 классе и 34 часов в 11 классе.

Учебный курс относится к циклу дисциплин по выбору.

Целью освоения курса «Основы анализа данных» является развитие практических навыков в сфере анализа данных и машинного обучения. В процессе изучения курса обучающиеся познакомятся с современными библиотеками машинного обучения и обработки данных: Pandas, Numpy, Seaborn, SciPy и Pytorch, и научатся применять их для решения задач анализа данных, машинного и глубинного обучения. В рамках курса учащимся будут представлены примеры этапов решения задачи анализа данных таких как: предобработка данных, разведывательный анализ, выдвижение гипотез и их проверка, выбор и обучение модели машинного обучения. Также обучающиеся получат необходимые навыки для участия в олимпиадах по анализу данных и искусственному интеллекту.

1. **Содержание учебного предмета**

**10 класс.**

**Тема 1. Знакомство с Python.**

Базовые понятия языка: переменные, условия, циклы. Функции, методы и операторы.

Библиотеки Pandas, Numpy. Работа с датасетами, манипуляции с данными и получение базовых статистик.

Типы данных, создание новых признаков

**Тема 2. Основы теории вероятности и статистики**.

Генеральная совокупность и выборка: частотные таблицы и распределения

Матожидание и разброс

Z-оценка, корреляция, выбросы

Тестирование гипотез, непараметрический критерии

Линейная регрессия и логистическая регрессия

**Тема 3. Визуализация данных**

Работа с библиотеками Python для визуализации данных: Matplolip, Seaborn, Plotly.

Изучение видов графиков: столбчатая диаграмма, гистограмма, пузырьковая диаграмма, box plot, круговая диаграмма. Чтение графиков и получение выводов из них.

Визуализация гипотез и анализа данных для представления на публике.

**Тема 4. Основы математики машинного обучения.**

Линейная алгебра. Матрицы, векторы, скалярное произведение, линейная независимость, векторные пространства, определитель матрицы, обращение матрицы, решение системы линейных уравнений, метод наименьших квадратов, матричные разложения.

Математический анализ. Дифференциал и производная, дифференцирование функций, дифференцирование в частных производных, градиент функции.

**Тема 5. Машинное обучение.**

Знакомство с библиотекой SciPy.

Задачи классификации, регрессии, ранжирования, кластеризации.

Линейная регрессия, резуляризация, логистическая регрессия, метод опорных векторов, многоклассовая классификация, решающие дервья, градиентный бустинг, обучение ранжированию, рекомендательные системы.

**Тема 5. Основы глубинного обучения**

Автоматическое дифференцирование, полносвязные нейронные сети.

Оптимизаторы и принципы их работы. SGD, Adam, AdamW, Dropout, Нормализация

Архитектура сверхточных нейронных сетей.

Задачи компьютерного зрения: классификация, сегментация, детекция.

**3. Тематическое планирование**

**10 класс**

|  |  |  |  |
| --- | --- | --- | --- |
| **№ п/п** | **Тематическое содержание занятия** | **Обязательные часы** | **Основные виды деятельности** |
| 1 | Знакомство с Python | 6 | Лекция и дискуссия по обозначаемой преподавателем проблематике. Решение задач. Выполнение практических заданий на компьютере. |
| 2 | Основы теории вероятности и статистики | 18 | Лекция и дискуссия по обозначаемой преподавателем проблематике. Решение задач. Выполнение практических заданий на компьютере. |
| 3 | Основы математики машинного обучения | 20 | Лекция и дискуссия по обозначаемой преподавателем проблематике. Решение задач. Выполнение практических заданий на компьютере. |
| 4 | Машинное обучение | 24 | Лекция и дискуссия по обозначаемой преподавателем проблематике. Решение задач. Выполнение практических заданий на компьютере. |
|  | **Всего** | **68** |  |

**11 класс**

|  |  |  |  |
| --- | --- | --- | --- |
| **№ п/п** | **Тематическое содержание занятия** | **Обязательные часы** | **Основные виды деятельности** |
| 1 | Машинное обучение  | 10 | Лекция и дискуссия по обозначаемой преподавателем проблематике. Решение задач. Выполнение практических заданий на компьютере. |
| 2 | Основы глубинного обучения | 22 | Лекция и дискуссия по обозначаемой преподавателем проблематике. Решение задач. Выполнение практических заданий на компьютере. |
| 5 | Резерв | 2 |  |
|  | **Всего** | **34** |  |

**Дополнительная литература**

**Учебно-методическое обеспечение образовательной деятельности**

1. Поляков К.Ю., Еремин, Е.А. Информатика. Углубленный уровень. Учебник для 11 класса. В 2-х частях. Часть 1. М.: БИНОМ Лаборатория знаний.

2. Поляков К.Ю., Еремин, Е.А. Информатика. Углубленный уровень. Учебник для 11 класса. В 2-х частях. Часть 2. М.: БИНОМ Лаборатория знаний.

3. Marc Peter Deisenroth. Mathematics for Machine Learning (2020). Cambridge University Press; 1st edition

4. Флах, П., Машинное обучение : наука и искусство построения алгоритмов, которые извлекают знания из данных, 2015

5. Шень А., Вероятность: Примеры и задачи. Издательство МЦНМО, 2016

6. Син Такахаси, Иноуэ Ирона. Регрессионный анализ, Додека XXI век, 2010.