



**КОНКУРСНОЕ ЗАДАНИЕ
ПЕРВОГО ЭТАПА ОТБОРА УЧАСТНИКОВ
НА УРАЛЬСКУЮ ПРОЕКТНУЮ СМЕНУ В СИРИУСЕ 2023**

Инструкция по выполнению конкурсного задания

1. **Задание единое** – одно для всех, состоит из трёх блоков: естественно-научного, инженерного и блока информационных технологий.
2. В каждом блоке задания разбиты на **2 уровня сложности**: I уровень (базовый), II уровень (продвинутый).
3. Решать все задания – **необязательно!** Каждый участник выполняет столько заданий, сколько может выполнить.
4. Если у вас возникнет вопрос по заданию, пожалуйста, напишите его на нашу почту: schooltalents@urfu.ru

Общие требования к оформлению задания

Выполненное задание оформляется в Microsoft Word и должно иметь следующие обязательные элементы и разделы:

1. **Титульный лист** с идентификацией участника: ФИО, школа, класс.
2. Пожалуйста, **обязательно указывайте блок и номер** выполняемого задания: естественно-научный, инженерный и блок информационных технологий.
3. Каждое задание должно начинаться с заголовка «**Блок _____, уровень _____, задание _____**».
4. Нумерация страниц **обязательна**.
5. Все дополнительные материалы или задания, требующие другой формат, прилагаются **ссылками внутри файла с выполненным заданием**.
6. Файл с выполненным заданием назвать: «**Фамилия_Имя_Отчество**» участника.

Критерии оценки заданий первого этапа:

- методика выполнения конкурсного задания;
- творческий подход к выполнению конкурсного задания;
- качество результата.

Самостоятельность работы над Конкурсным заданием является обязательным требованием. Экспертный совет в праве не засчитывать баллы за конкретное задание, в процессе выполнения которого не прослеживается личный вклад участника.

Желаем удачи!

ЕСТЕСТВЕННО-НАУЧНЫЙ БЛОК

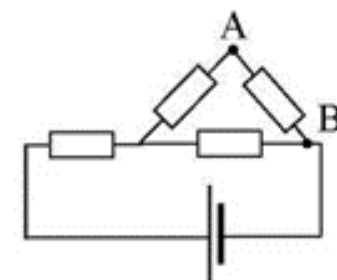
(за выполнение всех заданий в данном блоке можно набрать до 30 баллов)

I уровень - базовый

(за выполнение заданий данного уровня можно получить до 10 баллов)

Задание 1 (до 5 баллов)

У ученика имеются идеальная батарейка и резисторы с номиналами сопротивлений $R_1=4$ Ом, $R_2=0$ Ом, $R_3=2$ Ом, $R_4=12$ Ом. На представленной электрической схеме расположите резисторы таким образом, чтобы напряжение между точками А и В было максимально возможным.



В качестве ответа и укажите на схему номера резисторов или значения сопротивлений. Ответ обоснуйте.

Задание 2 (до 5 баллов)

Органические вещества, особенно природные, имеют сложную структуру и анализируются целым комплексом методов. Одним из базовых является элементный анализ – определение наличия и массового содержания определенного набора элементов, входящих в состав пробы. Типичными элементами-органогенами являются С, Н, N, O, S, P, Cl, щелочные и щелочноземельные металлы. Определить все элементы сразу нет возможности, поэтому определяют, обычно, основную тройку – С, Н, N.

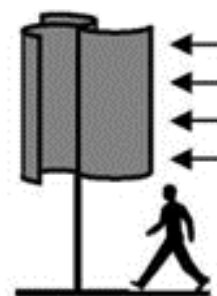
Приведите блок-схему прибора для СНN-анализа. Какие химические реакции лежат в его основе. Какие элементы являются мешающими для основной тройки и почему, как устраняют их негативный эффект? Какие особенности появляются при наличии ионов металлов в образце? Каким способом можно количественно определить содержание элементов?

II уровень – продвинутый

(за выполнение данного задания можно получить до 20 баллов)

Задание 3 (до 10 баллов)

Электрическая ветровая установка вырабатывает электроэнергию в условиях, заданных на рисунке. Необходимо оценить электрическую мощность данной установки? Для получения численного ответа, предлагается самостоятельно задать недостающие величины и выбрать их числовые значения.



Решение должно быть развёрнутым и обоснованным.

Задание 4 (до 10 баллов)

Из школьной программы все хорошо знают, что более активный металл (левее в электрохимическом ряду) вытесняет менее активный из растворов его солей (реакция замещения). Но есть способы обойти это ограничение, причем некоторые имеют высокую практическую значимость.

Обоснуйте теоретически возможность нарушения школьного правила, приведите два примера, когда металл, стоящий правее в стандартном электрохимическом ряду, растворяется под действием катиона, стоящего правее (включая водород в виде HCl). Где применяются такие «перевернутые» реакции?

ИНЖЕНЕРНЫЙ БЛОК

(за выполнение всех заданий в данном блоке можно набрать до 30 баллов)

I уровень - базовый

(за выполнение заданий этого уровня можно получить до 10 баллов)

Задание 1 (до 5 баллов)

Напишите в ряд n чисел «2» и поставьте между каждой парой соседних чисел один из знаков «+», «-», «×», «:», если хотите, добавьте несколько пар скобок так, чтобы получилось выражение, значение которого равно 2017. Используйте как можно меньше двоек: оценка за задачу определяется формулой $21 - n$, где n – количество двоек, которое Вы использовали.

Предоставьте пример, содержащий n двоек, с расстановкой знаков и скобок между любыми двумя соседними двойками.

Задание 2 (до 5 баллов)

Исполнитель Черепаха действует на плоскости с декартовой системой координат. В начальный момент Черепаха находится в начале координат, её голова направлена вдоль положительного направления оси ординат, хвост опущен. При опущенном хвосте Черепаха оставляет на поле след в виде линии. В каждый конкретный момент известно положение исполнителя и направление его движения.

У исполнителя существует две команды: **Вперёд n** (где n – целое число), вызывающая передвижение Черепахи на n единиц в том направлении, куда указывает её голова, и **Направо m** (где m – целое число), вызывающая изменение направления движения на m градусов по часовой стрелке.

Запись **Повтори k [Команда1 Команда2 ... КомандаS]** означает, что последовательность из S команд повторится k раз.

Черепахе был дан для исполнения следующий алгоритм:

Повтори N [Направо 25 Вперёд 5 Направо 25]

- 1) Найдите минимальное значение числа N , при котором Черепаха оставит след в виде замкнутой ломаной линии.
- 2) Найти минимальную площадь прямоугольника, включающего данную замкнутую ломаную линию.

II уровень – продвинутый

(за выполнение заданий данного уровня можно получить до 20 баллов)

Задание 3 (до 10 баллов)

Миша начертил на доске треугольник ABC такой, что $AB - BC = a$, a – некоторое действительное число. Затем он отметил точки M и N – середины сторон AB и AC соответственно, после чего стёр с доски всё, кроме точек M и N . С помощью циркуля и линейки восстановите треугольник ABC .

Предоставьте подробное решение с построением, доказательством того, что полученный треугольник – искомый, исследованием количества решений в зависимости от параметра a .

Задание 4 (до 10 баллов)

Квадрат разлинован на 30×30 клеток, в каждой клетке которого записано целое число, см. рисунок к заданию. Квадрат ограничен внешними стенами. Между соседними клетками квадрата также могут быть внутренние стены. Исполнитель Робот может перемещаться по клеткам, выполняя за одно перемещение одну из двух команд: **вверх** или **влево** на любое количество клеток в пределах квадрата и стен. Все стены отмечены на рисунке. Сквозь стену Робот пройти не может. Посетив клетку, Роботу дают столько монет сколько указано в данной клетке.

Определите максимальную и минимальную денежные суммы, которые может собрать Робот, пройдя из правой нижней клетки в верхнюю левую.

2	12	2	3	4	1	3	5	12	6	7	21	4	56	21	4	45	6	4	3
13	23	21	2	3	4	5	32	4	24	3	53	4	5	12	10	23	43	2	5
9	7	17	2	5	0	11	5	2	5	14	59	26	2	5	0	11	5	2	5
5	3	13	6	35	10	3	6	6	7	10	3	13	6	35	10	3	6	5	6
1	2	19	2	8	6	2	5	2	29	6	2	5	2	8	6	2	5	2	5
44	5	13	2	5	2	8	10	10	2	5	0	11	5	2	5	0	11	5	2
2	5	0	11	5	10	3	9	6	6	35	10	3	6	6	35	10	3	6	8
6	35	10	3	6	6	2	5	2	2	8	6	2	5	2	8	6	2	5	2
2	5	0	11	5	25	8	1	8	23	8	6	7	8	7	5	1	8	6	8
6	35	10	3	6	7	9	5	24	33	9	5	1	3	6	20	9	2	21	10
2	8	6	2	5	32	3	50	6	34	22	7	15	20	7	17	5	14	8	33
2	8	6	2	5	2	23	49	5	11	5	14	51	5	31	52	24	3	32	11
4	12	8	42	4	14	31	2	1	5	37	4	8	35	100	10	8	13	16	18
20	6	41	6	0	0	25	14	23	11	7	13	6	11	23	4	55	3	6	1
0	10	1	36	5	4	22	44	5	26	3	32	17	9	30	11	9	18	5	7
2	0	35	34	2	6	8	10	11	16	47	5	45	45	25	16	32	9	29	10
4	12	5	12	10	4	5	19	5	31	44	23	27	21	6	5	38	3	31	20
31	4	22	2	43	4	1	30	2	1	13	30	10	0	7	11	13	32	21	39
9	6	46	4	7	46	22	3	1	30	9	13	11	28	3	12	19	10	6	5
115	30	34	15	37	17	11	50	22	20	6	5	3	43	22	29	0	5	10	40

Рисунок к заданию №4

БЛОК ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ

(за выполнение всех заданий в данном блоке можно набрать до 30 баллов)

Блок информационных технологий состоит из заданий по 6 популярным IT-областями, а именно: UI/UX-дизайн, Frontend-разработка, Backend-разработка, Мобильная разработка, ML & Data science. Каждая область представлена двумя заданиями - базового и продвинутого уровней. Решение хотя бы одного задания базового уровня является достаточным подтверждением компетенций в соответствующей IT-области. Если есть желание попасть в определенное направление, то мы рекомендуем сфокусироваться на двух заданиях одной IT-области, и постараться решить их максимально качественно.

Да пребудет с вами Google!

UI/UX Дизайн

(за выполнение заданий до 6 баллов)

I уровень - базовый

(за выполнение задания данного уровня можно получить до 2 баллов)

В ходе выполнения данной задачи вам предстоит спроектировать мобильное приложение, помогающее правильно организовывать учебное время, в том числе кружки и секции, содержащее в себе ряд функций, полезных потенциальным пользователям. Для выполнения задачи, в первую очередь, следует провести базовое исследование пользователей и их потребностей с целью сформулировать ряд гипотез по полезным функциям проектируемого приложения. Именно это и требуется сделать на базовом уровне.

Задача

Провести базовое исследование потенциальных пользователей приложения-помощника в планировании учебного времени, с целью формирования ряда продуктовых гипотез.

Цель

Сформировать и приоритизировать список функций, которые могут быть максимально полезны пользователю.

Как решать?

Представьте, что вы делаете приложение-помощник для контроля и организации учебного времени и успех вашего проекта зависит от того, насколько простым, удобным и функциональным вы его сделаете (так оно в жизни, от части, и работает). Но что оно должно уметь? Как должно быть устроено? Какие функции будут использовать ваши пользователи, и в каких ситуациях? Ваша задача – это предположить и обосновать свои предположения. Выделите сегмент пользователей, для которого вы делаете приложение, опишите его в виде одного конкретного «персонажа» и выделите его «боль» (ключевую проблему, которую вы решите своим приложением). Т.е. вам нужно описать какой человек и на какую работу в своей жизни нанимает ваш продукт. Эта методика называется JTBD (JobsToBeDone). Рекомендуем с ней ознакомиться, чтобы выполнить задание максимально хорошо. Продумайте до 3-х таких «персонажей» (один сегмент пользователей - один «персонаж»).

Пример

Арина – девушка, 22-30 лет, работает 8 часов в сутки, нет мужа и детей. Арина проходит обучение в вузе, и посещает дизайнерские курсы и мастер классы. Однако, она часто не успевает посетить лекцию или мастеркласс. Мы можем помочь ей, предложив функцию уведомления в то время, а также проложить маршрут онлайн до места проведения мастеркласса. Таким образом, «Часто не успевает прослушать лекцию или посетить мастеркласс» – это «боль». «Функция уведомления в то время, когда требуется посетить занятие» – это гипотеза, касательно того, что такая функция поможет решить «боль» Арины. Пропишите персонажей (не более 3-х) и для каждого из них составьте таблицу с двумя столбцами («боли» и «гипотезы»).

Важно

Оценивается не количество гипотез (предполагаемых функций приложения), а качество и целесообразность. Не стесняйтесь опрашивать знакомых, похожих на ваших «персонажей», чтобы выделить какие из функций нужны/не нужны, важнее/не важнее остальных. Правильные вопросы для таких интервью поможет составить понимание методики Customer Interview. В рамках данной задачи владение этой методикой не требуется, но в дальнейшем будет очень полезным.

Примечание

Решение необходимо предоставить в виде ссылки на проект в Figma с возможностью редактирования («Share» → «Anyone with the link» → «Can edit») или в виде ссылки на презентацию в формате PDF. Ваше решение обязательно должно содержать описание от одного до трёх «персонажей» и таблицу «боли»/«гипотезы» для каждого из них. В таблице должно быть минимум по 4 «боли», каждой из которой соответствует своя «гипотеза».

Важно! Проверьте доступность ваших ссылок через режим инкогнито в браузере.

II уровень – продвинутый

(за выполнение задания данного уровня можно получить до 4 баллов)

В ходе выполнения задания базового уровня мы определили, кто является нашими целевым пользователями, и даже четко описали этих людей. Мы догадываемся об их проблемах, и о том, как решить их с помощью нашего приложения. Настало время воплотить его в жизнь (пока только со стороны дизайна).

Задача

Создать ряд макетов экранов предполагаемого приложения в Figma, содержащих в себе реализацию самых приоритетных функций из предыдущего задания базового уровня.

Требования к результату

- проект в Figma содержит в себе минимум 5 макетов экранов (размер фрейма: 375 x 812) - 1 балл;
- решение запускается как прототип в figma (в прототипе задействованы все макеты (экраны), присутствует правильный подход к навигации, понятна информационная архитектура) - 1 балл;
- спроектировано и отражено в прототипе как минимум 4 функции из отобранных в предыдущем задании 1 уровня - 1 балл;

Целостность и логическая завершенность спроектированного интерфейса будет большим плюсом. Важно, чтобы логических дыр в процессе использования приложения было как можно меньше.

Пример

<https://www.figma.com/file/upHDOPVNeBTfMS5b370xhk/?node-id=0%3A1>

Примечание

Решение необходимо предоставить в виде ссылки на проект в Figma с возможностью редактирования («Share» → «Anyone with the link» → «Can edit»).

Важно! Проверьте доступность ваших ссылок через режим инкогнито в браузере.

Перед отправкой проекта на проверку, дайте ваш figma-прототип в руки кому-то, кто, потенциально, может быть пользователем такого решения (Например, ваша подруга или друг, брат или сестра, которым нужно контролировать обучение). Попросите их в определенном порядке воспользоваться спроектированными функциями, и вы сразу увидите, что и где сделано неудобно, непонятно, или просто не нужно.

Frontend-разработка

(за выполнение заданий до 5 баллов)

I уровень - базовый

(за выполнение задания данного уровня можно получить до 2 баллов)

В этом задании необходимо реализовать веб-страницу, отображающую вход и регистрацию на одной странице, используем:

- React
- Vue

Требования к заданию:

- сверстать экран «Front 1», согласно дизайну в Figma-макете (страница «Front I»);
- все элементы должны быть представлены разными компонентами;
- страница должна быть адаптивной;
- использование CSS-препроцессоров или CSS-in-JS библиотек (tailwind, emotion, styled) будет плюсом;
- реализация анимации полос графика будет плюсом.

Материалы задания

Figma-макет:

<https://www.figma.com/file/KsW3I1NGJyxOiHTAQkPUW2/frontend?node-id=4%3A2>

Что посмотреть?

React Getting started (ru): <https://ru.reactjs.org/>

Vue Getting started: <https://vuejs.org/v2/guide/>

Примечание

Для демонстрации выполненной работы запишите видео с результатом работы и продемонстрируйте адаптивность страницы. Видео загрузите на YouTube и прикрепите ссылку на него в документе с ответами на задания. Исходный код загрузите на GitHub и прикрепите ссылку на репозиторий в ваше решение. Позаботьтесь о защите своего кода от нахождения другими участниками отбора.

Важно! Проверьте доступность ваших ссылок через режим инкогнито в браузере.

II уровень – продвинутый

(за выполнение задания данного уровня можно получить до 4 баллов)

Реализуйте дополнительную страницу личного кабинета.

Для этого нужно решить следующие задачи:

1. Сверстать макеты экранов по предложенному дизайну (вход/регистрация, личный кабинет). Дизайн экранов находится на странице «Front 2» в Figma-макете. При создании должен использоваться любой CSS-препроцессор или CSS-in-JS библиотека.
2. Результат работы должен представлять из себя рабочее веб-приложение, в котором можно пройти флоу от логина до экрана личного кабинета.

Бизнес-логику (работу с API, оплату) реализовывать не нужно: данные лучше хардкодить.

Большим плюсом будет демонстрация работы своего приложения, развернутого на любом бесплатном frontend-only хостинге (Например, Vercel или Netlify).

Примечание

Для демонстрации выполненной работы запишите видео с результатом работы и продемонстрируйте адаптивность страницы. Видео загрузите на YouTube и прикрепите ссылку на него в документе с ответами на задания. Исходный код загрузите на GitHub и прикрепите ссылку на репозиторий в ваше решение. Позаботьтесь о защите своего кода от нахождения другими участниками отбора.

Важно! Проверьте доступность ваших ссылок через режим инкогнито в браузере.

Backend-разработка

(за выполнение заданий до 6 баллов)

Допускается решение задания на следующих фреймворках/технологиях: Node.JS (Express / Nest.JS), Python (Django / Flask / FastAPI), Go, Rust, C# (ASP.NET + EF). В качестве базы данных допускается использование: MongoDB, Postgresql, SQLite.

I уровень - базовый

(за выполнение задания данного уровня можно получить до 2 баллов)

Реализуйте простой REST API с использованием любого вышеприведенного фреймворка.

Описание задачи: <https://pastebin.com/QADa0qpF>

Что посмотреть?

<https://metanit.com/web/nodejs/1.1.php>

<https://www.djangoproject.com/start/>

<https://fastapi.tiangolo.com/tutorial/>

<https://semaphoreci.com/community/tutorials/building-and-testing-a-rest-api-in-go-with-gorilla-mux-and-postgresql>

II уровень – продвинутый

(за выполнение задания данного уровня можно получить до 4 баллов)

Необходимо реализовать упрощённое GraphQL API библиотеки.

Требования к результату

Описание схемы данных:

Книга (Book)

- ID
- Название
- Максимальный период аренды
- Ссылка на обложку
- Автор (Author)

Автор (Author)

- ID
- Имя
- Ссылка на фотографию
- Дата рождения
- Дата смерти (если мёртв)
- Книги (Book)
- Количество доступных для аренды книг в хранилище библиотеки

Читатель (Reader)

1. ID
2. Имя
3. Почта
4. Аренды (Rent)
5. Штраф за просрочки

Аренда (Rent)

1. ID
2. Книга (Book)
3. Читатель (Reader)
4. Время аренды
5. Период, на сколько книга арендована
6. Размер потенциального штрафа за каждый день просрочки

Для моделей книг, читателей и авторов обязательно нужно реализовать GraphQL методы для CRUD операций (Создание, чтение, редактирование, удаление - Create, Read, Update, Delete).

Методы, которые нужно реализовать в API, помимо вышеприведенных CRUD-методов для основных моделей:

- Queries:
 - books - получить все книги в сервисе;
 - authors - получить всех авторов;
 - rents (readerId: ID!) - получить все аренды читателя.
- Mutations:
 - rentBook (bookId: ID!, readerId: ID!) – взять книгу в аренду.
 - returnBook (bookId: ID!, readerId: ID!) – вернуть книгу.

При сдаче книги, если сдача просрочена, клиенту должен быть начислен штраф за каждый день просрочки, согласно указанной цене.

Большим плюсом будет демонстрация работы своего API, развернутого на любом бесплатном хостинге (например: Heroku).

Примечание

Для демонстрации выполненной работы запишите видео, на котором вы делаете graphql-запросы к API через интерфейс GraphQL Playground. Видео загрузите на YouTube и прикрепите ссылку на него в документе с ответами на задания. Исходный код загрузите на GitHub и прикрепите ссылку на репозиторий в ваше решение. Позаботьтесь о защите своего кода от нахождения другими участниками отбора.

Важно! Проверьте доступность ваших ссылок через режим инкогнито в браузере.

Что посмотреть?

Getting Started (ru) - <https://webdevblog.ru/vvedenie-v-graphql/>

https://www.youtube.com/watch?v=F4vHSHzpO1g&ab_channel=HolyJS

Список публичных GraphQL API - <https://github.com/APIs-guru/graphql-apis>

Apollo Server - <https://www.apollographql.com/docs/apollo-server/getting-started/>

Мобильная разработка

(за выполнение заданий до 6 баллов)

I уровень - базовый

(за выполнение задания данного уровня можно получить до 2 баллов)

Реализуйте мобильное приложение, отображающее страницу / платформ: React Native (not expo), Flutter, Native Android (Kotlin, Java), Native iOS (Swift), Native Android / iOS + KMP.

В требования задачи входит:

7. Произвести установку и настройку среды для разработки мобильных приложений (Android Studio / Xcode).
8. Сверстать экран "Промокод", согласно дизайну в Figma-макете (страница "Mob I").
9. Все элементы графика должны быть представлены разными компонентами.
10. Страница должна быть адаптивной.

Материалы задания

Figma-макет:

<https://www.figma.com/file/t7yZ15UOqEkfBjGcSE6sjx/Untitled?node-id=0%3A1>

Что посмотреть?

Getting started (React Native): <https://riptutorial.com/ru/react-native>

Getting started (Flutter): <https://flutter.dev/docs/get-started/codelab>

Примечание

Для демонстрации выполненной работы запишите видео с результатом работы и продемонстрируйте адаптивность страницы. Видео загрузите на YouTube и прикрепите ссылку на него в документе с ответами на задания. Исходный код загрузите на GitHub и прикрепите ссылку на репозиторий в ваш файл с ответами. Позаботьтесь о защите своего кода от нахождения другими участниками отбора.

Важно! Проверьте доступность ваших ссылок через режим инкогнито в браузере.

II уровень – продвинутый

(за выполнение задания данного уровня можно получить до 4 баллов)

Реализуйте часть мобильного приложения для входа, регистрации и восстановления аккаунта.

Выполните следующие задачи:

6. Сверстать макеты экранов по предложенному дизайну. Дизайн экранов находится в Figma-макете.
7. Результат работы должен представлять из себя рабочее мобильное приложение, в котором можно пройти флоу от login до экрана восстановления аккаунта.

Бизнес-логику (регистрацию, вход, восстановление) реализовывать не нужно: данные лучше хардкодить.

Примечание

Для демонстрации выполненной работы запишите видео с результатом работы и продемонстрируйте работу приложения. Видео загрузите на YouTube и прикрепите ссылку на него в документе с ответами на задания. Исходный код загрузите на GitHub и прикрепите ссылку на репозиторий в ваш файл с ответами. Позаботьтесь о защите своего кода от нахождения другими участниками отбора.

Важно! Проверьте доступность ваших ссылок через режим инкогнито в браузере.

Материалы задания

Figma-макет: <https://www.figma.com/file/t7yZ15UOqEkfBjGcSE6sjx/Untitled?node-id=0%3A1>

Data science & Machine learning

(за выполнение заданий до 6 баллов)

I уровень - базовый

(за выполнение задания данного уровня можно получить до 2 баллов)

Выполните следующее задание:

Напишите парсер данных пользователей вк, другими словами - функцию/класс/скрипт, который принимает на вход ссылку на профиль пользователя ВКонтакте и выводит на экран:

- Имя пользователя
- Возраст
- Статус
- Названия групп (сообществ), на которые подписан пользователь и ссылки на эти группы
- Количество друзей пользователя
- Количество подписчиков
- Среднее количество лайков на посте

Для случайного набора из 10000 пользователей ВКонтакте (чтобы собрать такую можно, например, сгенерировать список рандомных id) запустите написанный на предыдущем шаге парсер, на этот раз записывая скачиваемые данные в csv файл (см. pandas.to_csv()) CSV файл должен иметь следующие столбцы(см. таблицу 1)

Требования к оформлению работы:

- Скрипты из шага 1 и шага 2 в виде ссылки на код/файлов/архива с файлами и тд
- Итоговый csv файл

Материалы:

- Как парсить данные:
<https://dev.vk.com/method/users.get>
<https://dev.vk.com/groups.get>

...

(всю информацию о парсинге конкретных данных можно найти в документации вк (выше) или гугле)

- Как сформировать csv файл
https://pandas.pydata.org/docs/reference/api/pandas.DataFrame.to_csv.html

Да и в целом pandas - очень удобная и часто используемая библиотека

Таблица 1.

Название столбца	Описание	Тип данных в столбце
ID	ID пользователя	число/строка
name	Имя пользователя	строка
age	Возраст	число
status	Статус пользователя	строка
groups	Список названий групп, на которые подписан пользователь (можно писать в таблицу список одной строкой, разделяя названия групп знаком "\n")	строка
groups_links	Список ссылок на группы, на которые подписан пользователь (можно писать в таблицу список одной строкой, разделяя ссылки знаком "\n")	строка
friends	Количество друзей	число
followers	Количество подписчиков	число
likes	Среднее количество лайков на посте	число
relationship	Семейное положение	строка

II уровень – продвинутый

(за выполнение задания данного уровня можно получить до 4 баллов)

Выполните следующее задание:

Обучите модель машинного обучения, которая, основываясь на странице пользователя ВКонтакте, предсказывает его возраст. Для этого

- Разделите таблицу выше на тестовый и тренировочный наборы данных (в тестовую часть должно попасть 20% выборки, в обучающую (тренировочную) - 80%).
- Выделите колонку “age” в отдельную переменную, удалите из общей таблички (аналогично в приложенном tutoriale удаляется колонка ‘Survived’), её мы будем предсказывать
- Удалите колонки name, groups_links и ID из таблицы с данными
- На тренировочных данных обучите модель градиентного бустинга CatBoost, передав при обучении в качестве параметров cat_features=[‘relationship’] и text_features=[‘status’, ‘groups’] (CatBoost умеет сам обрабатывать категориальные и тестовые признаки)
- Замерьте качество итоговой модели на отложенных на шаге 1 тестовых данных, предварительно проведя с этой частью данных шаги 2-3

В данном задании также поощряется фантазия в вопросе выбора данных и подхода к ним: если хочешь - можешь подтягивать и использовать абсолютно любую информацию о пользователе, кроме даты его рождения и возраста. Например, хорошей идеей будет в качестве одной из фичей передавать средний возраст друзей пользователя или количество лет, прошедшее после указанного в профиле года окончания вуза.

Требования к оформлению работы

- Ссылка на код/ноутбук (colab notebook) или же сам файл кода/ноутбука
- Качество на тренировочной выборке
- Качество на тестовой выборке

Материалы:

Разделение выборки на тестовую и обучающую:

https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.train_test_split.html

Тьюториал CatBoost:

https://github.com/catboost/tutorials/blob/master/python_tutorial.ipynb

Какие бывают фичи:

<https://proglab.io/p/postroenie-i-otbor-priznakov-chast-1-feature-engineering-2021-09-15>

Обзор с примерами, что такое входные данные, фичи, как работать с библиотекой pandas и как улучшить качество подаваемых на вход модели данных.

<https://mlbootcamp.ru/ru/article/tutorial/>

Очень хорошим языком написанная теория по машинному обучению вместе с ключевыми терминами:

<https://ml-handbook.ru/>

Colab Notebook (онлайн среда разработки, часто используемая в ml)

<https://colab.research.google.com/>

Про машинное обучение есть куча статей на Хабр/Medium/towardsdatascience, которые выпадают при поиске в гугле материалов по запросу “машинное обучение”, + миллион курсов вроде <https://ml.ai-academy.ru/>, проходи любой курс и читай как статьи - это поможет разобраться в теме

Если в целом хочешь более серьезно развиваться в направлении машинного обучения - то вот очень крутая и исчерпывающая подборка материалов:

<https://habr.com/ru/post/589765/>