

Демонстрационный вариант
заданий заключительного этапа (полуфинал)
по направлению «Математическое моделирование»

Категория участия: «Бакалавриат»
(для поступающих в магистратуру)

1. Полубесконечный горизонтальный пористый пласт имеет вид прямоугольного параллелепипеда, все грани которого (кроме левой) непроницаемы для жидкости. Пласт насыщен жидкостью с коэффициентом объемной сжимаемости $\beta=1 \cdot 10^{-10} \text{ Па}^{-1}$, давление которой во всех точках в начальный момент времени постоянно и равно $p_0 = 5 \text{ МПа}$. На левой границе пласта давление в момент времени $t = 0$ мгновенно уменьшается до $p_e = 1 \text{ МПа}$ и в дальнейшем поддерживается постоянным. Определить значение давления в момент времени $t = 100 \text{ с}$ в точке пласта, отстоящей на расстоянии $x=10 \text{ м}$ от его левой границы, если зависимость вектора скорости течения жидкости в пористом пласте \vec{V} от давления p описывается линейным законом фильтрации: $m\vec{V} = -\frac{k}{\mu} \text{grad}p$, где $m = 0.1$ и $k = 10^{-14} \text{ м}^2$ – коэффициенты пористости и проницаемости пласта, а $\mu = 10^{-3} \text{ (Па} \cdot \text{с)}$ – коэффициент динамической вязкости жидкости. Ответ выразить в МПа и округлить до сотых.

Ответ: 3.08

2. Модель формирования стратегии в системе управления запасами предприятия. Запасы создаются на любом предприятии и зачастую причинами их создания являются различия в темпах производства и потребительского спроса. Под запасами можно понимать такое количество материалов, ресурсов, с помощью которого ведется непрерывно выпуск товаров. Сущность управления запасами заключается в установлении баланса между объемами заказа и вновь прибывшей партией товара.

Постройте экономико-математическую модель снижения издержек управления текущими многономенклатурными материальными запасами, построенную по месячным данным четырех хранимых продуктов цеха жилищного хозяйства ООО «Домовенок», с целью определить условия и последовательность закупок с учетом ограниченности общей складской площади, устанавливая при этом оптимальные значения ряда параметров эффективности функционирования.

Предприятие имеет несколько цехов. Каждый цех имеет свои склады, которые специализированы для хранения сырья и готовой продукции. Для полноценного функционирования предприятие закупает некоторые виды продукции у сторонних поставщиков и хранит их на своих складах. Необходимо разработать экономико-математическую модель снижения издержек управления текущими многономенклатурными материальными запасами, построенную по месячным данным четырех основных видов запасов цеха жилищного хозяйства предприятия непрерывного цикла по оказанию жилищно-коммунальных услуг и ремонтных работ.

Основные предположения модели

- рассматривается задача управления несколькими видами запасов;
- складское пространство ограничено;

Всероссийская олимпиада студентов «Я – профессионал»

- интенсивность спроса на ресурс (количество единиц ресурса, потребляемых в единицу времени) – постоянная величина (константа);
- удельные затраты на хранение в единицу времени (затраты на хранение единицы ресурса в единицу времени) – постоянная величина (константа);
- затраты на оформление, связанные с размещением заказа – постоянная величина (константа);
- заказ размещается через целое число дней и пополняется мгновенно;
- дефицит отсутствует.

Требуется:

1. Ответить, как должна быть определена оптимальная стратегия управления запасами, т.е. определить объем заказа и период между заказами, при которых общие среднемесячные затраты были бы минимальными.
2. Оценить экономический эффект от предложенной Вами стратегии, сравнив ее с используемой на предприятии (таблица 5).

Данные для модели.

Число дней в месяце $n = 30$.

Общая площадь складского помещения составляет $S = 1650$ тонн в месяц.

Полезная площадь склада составляет $S - S_0 = 1500$ тонн в месяц. (возможны отклонения не более чем на 5% за счет ежедневного изменения запасов и использования площадей, выделенных под технологические нужды)

Основные постоянные затраты и переменные затраты на доставку и хранение 1 тонны запасов в 1 месяце представлены в таблицах 1-4:

Таблица 1

Элемент переменных затрат	Переменные затраты на хранение запасов (в месяц)			
	Продукт			
	Кровельная сталь	Задвижка чугунная СЧ 15; 20; 25;	Топливо для технологических целей	Пенополиуретан ППУ-110
	1	2	3	4
1. Хранение (руб./тонн)	2650	2700	2550	2480
2. Страхование (руб./тонн)	200	300	150	120

Таблица 2

Постоянные затраты на хранение запасов	
Постоянные затраты на хранение запасов, независящие от величины запасов	Значение затрат, тыс.руб./месяц
1. Износ склада	4
2. Плата сторонним организациям за пожарную и сторожевую охрану	3,5
3. Налоги на запасы (имущество)	3

Таблица 3

Затраты по доставке 1 тонны материала (тыс.руб.)				
Вид затрат	Кровельная сталь	Задвижка чугунная 30ч 906; 6бк	Топливо для технологических целей	Пенополиуретан ППУ-110
1. Доставка ж/д транспортом	2,1	2,2	1,7	1,8

Всероссийская олимпиада студентов «Я – профессионал»

2. Подача вагонов	0,75	0,75	1,3	0,75
3. Прочие затраты	1,15	1,25	1,12	1,05
4. Выгрузка-складирование	0,78	0,6	1,1	0,82
5. Доставка	1,4	1,5	1,28	1,05

Таблица 4

Основные параметры для заказа продукции за месяц

Продукция	Месячный объем потребления, тн.	Площадь для хранения продукта, (м ² на тонну)
Кровельная сталь	55	35
Задвижка чугунная	72	28
Топливо для технологических целей	23	15
Пенополиуретан ППУ-110	45	27

Таблица 5

Вариант управления запасами принятый на предприятии.

Продукция	Объем партии, тонн	Частота заказа в месяц	Продолжительность между заказами, день
Кровельная сталь	14	4	8
Задвижка чугунная	15	5	6
Топливо для технологических целей	8	3	10
Пенополиуретан ППУ-110	15	3	10

Ответ: 1) Кровельная сталь и задвижки чугунные требуется заказывать 4 раза в месяц через каждые 8 дней, топливо для технологических целей можно заказывать 2 раза в месяц по 11,5 тонн, а пенополиуретан ППУ-110 – 3 раза в месяц по 15 тонн.

2) Экономический эффект по результатам многономенклатурной модели для предприятия составит 3,92 тыс.руб. в месяц

3. Набор данных о ценах на дома в Бостоне включает в себя прогноз цены дома в тысячах долларов с учетом деталей дома и его окрестностей. Эти данные были собраны в 1978 году, каждая из 506 записей представляет собой совокупную информацию о 14 характеристиках домов из различных пригородов, расположенных в Бостоне. Требуется построить модель, предсказывающую денежную стоимость дома, расположенного в районе Бостона.

Качество предсказания будет измеряться при помощи следующей метрики:

$$MAPE = \frac{1}{n} \cdot \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i}$$

Баллы за задачу будут начисляться по формуле:

$$5 + 40 * (0.5 - MAPE),$$

если метрика $MAPE$ оказалась меньше 0.5; в противном случае за задачу дается 0 баллов.

Всероссийская олимпиада студентов «Я – профессионал»

Формат входных данных: обучающий датасет train.tsv находится в архиве, доступном по адресу: <https://disk.yandex.ru/d/TYuX9cJ0riiy7w?w=1>.

Файл содержит следующие поля:

CRIM: уровень преступности на душу населения по городам

ZN: доля жилой земли для участков более 25 000 кв. футов

INDUS: доля неторговых площадей в каждом городе

CHAS: фиктивная переменная реки Чарльз (она равна 1, если тракт ограничивает реку; 0 в противном случае)

NOX: концентрация оксидов азота (частей на 10 миллионов)

RM: среднее количество комнат на одно жилище

AGE: доля собственников владеющих собственностью, построенной до 1940 года

DIS: взвешенные расстояния до пяти бостонских центров занятости

RAD: показатель доступности радиальных магистралей

TAX: полная стоимость налога на недвижимость на 10 000 долларов

PTRATIO: соотношение учеников и учителей по городам

B: доля людей афроамериканского происхождения по городам, рассчитанная по специальной формуле

LSTAT: процент людей с низким социальным статусом.

PRICE: средняя стоимость домов, занимаемых владельцами, в тыс. долларов.

Формат вывода: Тестовый датасет test.csv состоит из $n + 1$ строк (считая первую строку — заголовок **PRICE**). Файл доступен по адресу: <https://disk.yandex.ru/d/TYuX9cJ0riiy7w?w=1>. Необходимо предоставить файл submission.csv, в котором в i -той строке будет записано значение цены в виде действительного числа для $(i + 1)$ -ой строки тестового датасета.

Ответ: Результаты решения находятся в файле answer.csv, доступном по тому же адресу.

4. Индейцы племени Пима из Аризоны и Мексики имеют самый высокий уровень распространенности диабета в мире. Был проведен анализ их медицинских карт, чтобы оценить, возможно ли предсказать начало диабета на основе диагностических мер.

Имеющийся набор данных взят из Национального института диабета, болезней органов пищеварения и почек. Цель набора данных – диагностически предсказать, есть ли у пациента диабет, на основе определенных диагностических измерений, включенных в набор данных. На выбор этих экземпляров из более крупной базы данных было наложено несколько ограничений. В частности, все пациенты здесь – женщины не моложе 21 года, принадлежащие к индейцам пима.

Требуется выполнить прогнозирование начала диабета в течение 5 лет у индейцев пима с учетом медицинских данных – более формально, необходимо предсказать значение параметра Outcome.

Качество классификатора будет измеряться при помощи метрики $f1$ — с правильно предсказанным параметром **Outcome** для всех людей внутри группы.

Баллы за задачу будут начисляться по формуле:

$$5 + 40 * (Q - 0.5),$$

где Q — метрика $f1$ на тестовом наборе данных, если она оказалась выше 0.5. В противном случае за задачу дается 0 баллов.

Формат входных данных: обучающий датасет train.tsv находится в архиве, доступном по адресу: <https://disk.yandex.ru/d/TYuX9cJ0riiy7w?w=1>.

Файл содержит следующие поля:

Pregnancies: число беременностей

Glucose: концентрация глюкозы в плазме через 2 часа при тесте

Blood pressure: диастолическое артериальное давление (мм рт.ст.)

SkinThickness толщина кожной складки трицепса (мм)

Insulin: 2-часовой сывороточный инсулин (м.е. / мл)

BMI: индекс массы тела (вес в кг / (рост в м)²)

DiabetesPedigreeFunction: генетическое влияние и наследственный риск

Age: возраст (лет)

Outcome: 1 наличие диабета, 0 отсутствие диабета

Формат вывода: Тестовый датасет test.csv состоит из $n + 1$ строк (считая первую строку — заголовок **Outcome**). Файл доступен по адресу: <https://disk.yandex.ru/d/TYuX9cJ0riiy7w?w=1>. Необходимо предоставить файл submission.csv, в котором в i -той строке будет записано значение предиктора (0 или 1) в виде действительного числа для $(i + 1)$ -ой строки тестового датасета.

Ответ: Результаты решения находятся в файле answer.csv, доступном по тому же адресу.

Спецификация для заключительного этапа (полуфинал)

Элемент спецификации	Комментарий к заполнению
Название направления	Математическое моделирование
Указание уровня подготовки	Бакалавриат
Описание целевой аудитории	<p>Данный комплект заданий подготовлен в рамках олимпиады «Я - профессионал» и предназначен для оценки знаний и навыков участников заключительного этапа олимпиады по профилю «Математическое моделирование». В качестве рекомендуемых направлений бакалавриата участников предлагаются следующие:</p> <p>01.03.01 Математика 01.03.02 Прикладная математика и информатика 01.03.03 Механика и математическое моделирование 01.03.04 Прикладная математика 02.03.01 Математика и компьютерные науки 02.03.03 Математическое обеспечение и администрирование информационных систем 09.03.03 Прикладная информатика</p>
Максимальное количество баллов	100 баллов
Время на выполнение	72 часа
Список ресурсов для самостоятельной подготовки	<ol style="list-style-type: none"> 1. Акопов, А. С. Имитационное моделирование. Учебник и практикум / А.С. Акопов. - М.: Юрайт, 2015. - 390 с. 2. Белов, П. Г. Управление рисками, системный анализ и моделирование. Учебник и практикум. В 3 частях. Часть 2 / П.Г. Белов. - М.: Юрайт, 2016. - 252 с. 3. Бродский, Г. Л. Экономико-математические методы и модели в логистике. Процедуры оптимизации / Г.Л. Бродский, Д.А. Гусев. - М.: Academia, 2012. - 288 с. 4. Введение в математическое моделирование. Учебное пособие. - М.: Логос, 2015. - 440 с. 5. Зайдель, А.Н. Математическое моделирование. Построение моделей и численная реализация: Учебное пособие / А.Н. Зайдель. - СПб.: Лань, 2016. - 304 с. 6. Математические модели систем управления. Учебное пособие. - М.: Издательство СПбГУ, 2000. - 340 с. 7. Моделирование систем / И.А. Елизаров и др. - М.: ТНТ, 2013. - 136 с. 8. Морозов, В.В. Исследование операций в задачах и упражнениях / В.В. Морозов, А.Г. Сухарев, В.В. Федоров. - Москва: Гостехиздат, 2016. - 595 с. 9. Павловский, Ю. Н. Компьютерное моделирование. Учебное пособие / Ю.Н. Павловский, Н.В. Белотелов, Ю.И. Бродский. - М.: Физматкнига, 2014. - 304 с. 10. Программирование, численные методы и математическое моделирование / И.Г. Семакин и др. - М.: КноРус, 2016. - 304 с. 11. Рейзлин, В. И. Математическое моделирование. Учебное пособие / В.И. Рейзлин. - М.: Юрайт, 2016. - 128 с. 12. Ржевский, С. В. Исследование операций. Учебное пособие / С.В.

	<p>Ржевский. - М.: Лань, 2013. - 480 с.</p> <p>13. Семакин, И.Г. Программирование, численные методы и математическое моделирование (для бакалавров) / И.Г. Семакин, О.Л. Русакова, Е.Л. Тарунин. - М.: КноРус, 2018. - 288 с.</p> <p>14. Степанов, В.И. Экономико-математическое моделирование / В.И. Степанов. - М.: Academia, 2018. - 336 с.</p> <p>15. Юдин, С. В. Математика и экономико-математические модели. Учебник / С.В. Юдин. - М.: Инфра-М, РИОР, 2016. - 376 с.</p> <p>16. Глубокое обучение (Часть I - Основы прикладной математики и машинного обучения) Goodfellow I. et al. Deep learning. –Cambridge :MIT press, 2016. – Т. 1.;</p> <p>17. Bishop C. M. Pattern recognition and machine learning. – Springer,2006.</p> <p>18. Введение в статистическое обучение Джеймс Г. и др. Введение в статистическое обучение с примерами на языке R / Пер. с англ. // СЭ Мастицкий. М.: ДМК-Пресс. – 2016.</p> <p>19. Николенко С. И., Кадури А. А., Архангельская Е. О. Глубокое обучение. – "Издательский дом "Питер", 2017.</p>
<p>Формат состязаний. Требования к содержанию и оформлению заданий</p>	<p>Подготовленный участником полуфинала ответ должен представлять собой электронный файл формата doc/docx/pdf, содержащий скан-копию (фотографии) решения задач, представленных на бумажном бланке. Электронный файл должен быть загружен в установленные сроки в личный кабинет участника Олимпиады.</p> <p>Требования к скан-копии (фотографии) ответа: высококонтрастное изображение с разрешением не менее 300 dpi.</p> <p>Допустимо представление решения задачи (части задачи) путем разработки программного продукта, но дополнительные баллы за применение данного способа решения не начисляются. В этом случае подготовленный участником полуфинала электронный отчет дополняется скриншотами работы программы с полученным решением задачи, а также листингом программы.</p> <p>По завершению полуфинала участнику даётся 10 минут на оцифровку рукописного решения, перевода его в указанный формат и загрузку в личном кабинете на интернет-портале Олимпиады.</p> <p>Для задачи машинного обучения в дополнение к имеющимся требованиям участнику необходимо представить:</p> <ol style="list-style-type: none"> 1. Программный код выполнения задачи с пояснениями основных моментов. 2. Ответ в виде файла формата csv с целью последующей загрузки и проверки решения на платформе Яндекс.контест.
<p>Дополнительная информация/инструкции для участников, которые не вошли в регламент по направлению</p>	<p>Для задачи машинного обучения в дополнение к имеющимся требованиям участнику необходимо представить:</p> <ol style="list-style-type: none"> 1. Программный код выполнения задачи с пояснениями основных моментов. 2. Ответ в виде файла формата csv с целью последующей загрузки и проверки решения на платформе Яндекс.контест.
<p>Краткое описание структуры задания и его</p>	<p>Задание состоит из четырех задач, раскрывающих блоки раздела «Математическое моделирование» – моделирование физических</p>

<p>основные характеристики. Система оценивания заданий.</p>	<p>процессов и процессов механики, моделирование социально-экономических процессов и процессов банковской сферы и методы решения задач математического моделирования (задача классификации и задача регрессии), основанные на машинном обучении. Каждое задание представляет собой постановку задачи практического применения. Система оценивания заданий подробно описана в пункте «Критерии оценивания».</p>
<p>Информация об элементах практикоориентированности в заданиях (участие работодателей в составлении заданий)</p>	<p>Все задания заключительного этапа носят практический характер и составлены совместно с генеральным партнером направления – акционерным обществом «Газпромбанк» и партнером по методической составляющей – научно-исследовательским и проектным институтом «РН-БашНИПИнефть». В рамках представленных заданий оцениваются навыки участников в построении математических моделей и их применении к решению задач социально-экономического и физического характера, а также навыки решения задач математического моделирования с применением технологий машинного обучения.</p>
<p>Критерии оценивания</p>	<p>Максимальный балл в случае успешного решения задач следующий: задача 1 – 25 баллов; задача 2 – 25 баллов; задача 3 – 25 баллов; задача 4 – 25 баллов.</p> <p>В случае частичного выполнения задачи критерии оценивания по задаче 1 (моделирование физических процессов и процессов механики) следующие: Итоговый балл участника K определяется по формуле</p> $K = 25 - K_{\text{ошибка отчет}} - K_{\text{численная погрешность}},$ <p>где $K_{\text{ошибка отчет}}$ – показатель снижения баллов за ошибки в отчете, принимающий значения из диапазона от 0 до 10: $K_{\text{ошибка отчет}} = 10$, если отчет не предоставлен; $K_{\text{ошибка отчет}} = 7 - 9$, если приведены ошибочные или фрагментарные выкладки $K_{\text{ошибка отчет}} = 4 - 6$, если выкладки в целом выполнены правильно, но конечный результат содержит ошибки. $K_{\text{ошибка отчет}} = 1 - 3$, если конечный результат в целом верен, но выкладки являются неполными или содержат незначительные недочеты, не влияющие на конечный результат $K_{\text{ошибка отчет}} = 0$, если результат верен, выкладки полные и безошибочные.</p> <p>$K_{\text{численная погрешность}}$ – показатель снижения баллов за погрешность отклонения решения S_j в точках j от точного аналитического решения S_j^*, принимающий значения: - 15, если txt-файл отсутствует либо численное решение неверно</p>

$$\left(\frac{\sum_{j=1}^{100} \frac{(S_j - S_j^*)}{S_j^*}}{100} \geq 1 \right);$$

$$- 15 \times \frac{\sum_{j=1}^{100} \frac{(S_j - S_j^*)}{S_j^*}}{100}, \text{ если предоставлен txt-файл и } \frac{\sum_{j=1}^{100} \frac{(S_j - S_j^*)}{S_j^*}}{100} < 1 .$$

Точное аналитическое решение S_j^* конкурсантам не предоставляется.

В случае частичного выполнения задачи критерии оценивания по задаче 2 (моделирование социально-экономических процессов) следующие (указан процент от максимального балла за задачу с округлением в сторону большего целого числа):

85-100% – Участник отчетливо демонстрирует понимание задачи, выполняет заданные требования, уверенно оперирует терминологией задачи, использует альтернативные варианты решения и аргументировано их обосновывает. Расчеты выполнены и оформлены верно.

51-84% – Участник демонстрирует понимание задачи, выполняет заданные требования, обоснование приводит краткое, но не ошибочное. Расчеты выполнены в целом правильно, но имеются незначительные недочеты, не влияющие на конечный результат.

21-50% – Участник в ответе затрагивает каждое из заданных требований, но задачу понимает не полностью, при формулировании решения или их обосновании демонстрирует неверное использование или неполное понимание терминологии и специфики задачи. Расчеты оформлены, но содержат ошибки.

1-20% – Участник демонстрирует минимальное понимание, не выполняет заданные требования, текст ответа содержит ошибки и неточности, не соответствует поставленной задаче. Расчеты практически полностью неверны.

0% – Ответ полностью неверен или отсутствует.

Критерий оценивания задачи 3 (задача регрессии): баллы будут определяться по формуле

$$\text{Балл} = 40 * (0.5 - \text{MAPE}) + R,$$

если метрика MAPE оказалась меньше 0.5; в противном случае за задачу дается 0 баллов. R – дополнительный балл от жюри (5 баллов) после оценки присланного решения.

Метрика MAPE вычисляется по формуле:

$$\text{MAPE} = \frac{1}{n} \cdot \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i}$$

здесь y_i – истинное значение, \hat{y}_i – прогнозируемое значение, n – количество предсказываемых значений.

Критерий оценивания задачи 4 (задача классификации): баллы будут

	<p>определяться по формуле</p> $\text{Балл} = 40 \cdot (Q - 0.5) + R$ <p>где Q – метрика f1 на тестовом наборе данных, если она оказалась выше 0.5. В противном случае за задачу ставится 0 баллов. R – дополнительный балл от жюри (5 баллов) после оценки присланного решения.</p> <p>Метрика f1 вычисляется по формуле:</p> $F1 = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$ $\text{Precision} = \frac{TP}{TP + FP}$ $\text{Recall} = \frac{TP}{TP + FN}$ <p>TP (True Positives) — верно классифицированные положительные примеры (так называемые истинно положительные случаи).</p> <p>TN (True Negatives) — верно классифицированные отрицательные примеры (истинно отрицательные случаи).</p> <p>FN (False Negatives) — положительные примеры, классифицированные как отрицательные (ошибка I рода). Это так называемый «ложный пропуск» — когда интересующее нас событие ошибочно не обнаруживается (ложно отрицательные примеры).</p> <p>FP (False Positives) — отрицательные примеры, классифицированные как положительные (ошибка II рода). Это ложное обнаружение, т.к. при отсутствии события ошибочно выносится решение о его присутствии (ложно положительные случаи).</p>
Наличие подробного примера решений демоверсии заданий	Нет (демоверсии заданий будут разобраны на вебинаре, предваряющем полуфинал).