

**Банк заданий по базовой части вступительного испытания в магистратуру**

**Задание 1 (5 баллов)**

**1.1. Вопрос:**

Имеются некоторые данные о переменных  $x$  и  $y$  (см. таблицу). Определить абсолютное значение коэффициента корреляции между переменными  $x$  и  $y$ . Является ли связь между этими переменными тесной или нет и почему? Данные можно округлить до сотых. Ответ должен содержать развёрнутое решение и пояснение.

X	12	22	11	13	15	14
Y	5	7	-4	3	1	1

**Ответ:**

Коэффициент детерминации  $R^2 = 0,43$

Абсолютное значение коэффициента корреляции:  $|r| = 0,66$ .

Интерпретация связи между переменными:

Коэффициент корреляции  $r = 0,66$  указывает на умеренную положительную линейную связь между переменными  $x$  и  $y$ .

Связь не является строго тесной (так как  $|r|$  не близок к 1), но она заметна и статистически значима для небольшого объема данных.

**1.2. Вопрос:**

Имеются некоторые данные о переменных  $x$  и  $y$  (см. таблицу). Используя F-статистику, оценить на уровне  $\gamma = 0,95$  значимость полученного уравнения регрессии. По критерию Стьюдента проверить значимость коэффициента линейной регрессии. Выполнить тест гипотез с использованием статистик Фишера, Стьюдента и их порогов. Данные можно округлить до сотых. Ответ должен содержать развёрнутое решение и пояснение.

y	6,8	6,2	5,9	5,7	5,5	5,3	4,3
x	4,1	5,9	5,2	6,8	5,8	4,2	5,2

**1.3. Вопрос:**

Имеются некоторые данные о переменных  $x$  и  $y$  (см. таблицу). Используя F-статистику, оценить на уровне  $\gamma = 0,95$  значимость полученного уравнения регрессии. По критерию Стьюдента проверить значимость коэффициента линейной регрессии. Данные можно округлить до сотых. Ответ должен содержать развёрнутое решение и пояснение.

y	6,8	6,2	5,9	5,7	5,5	5,3	4,3
x	5,1	5,9	7,2	1,8	8,8	7,2	5,2

**1.4. Вопрос:**

Имеются некоторые данные о переменных  $x$  и  $y$  (см. таблицу). Проверить гипотезу о значимости / незначимости коэффициентов уравнения регрессии на уровне  $\alpha = 0,05$  значимость полученного уравнения регрессии. По критерию Стьюдента проверить значимость коэффициента линейной регрессии. Данные можно округлить до сотых. Ответ должен содержать развёрнутое решение и пояснение.

y	14,2	14,4	15,1	15	13,9	16,9	12,6
x1	0,2	0,2	0,15	0,2	0,3	0,2	0,15
x2	2	1,2	3,3	4,3	5,3	2,3	9,9
x3	23,1	23,3	25,6	27,9	24,2	24,5	24,8

## Задание 2 (5 баллов)

### 2.1. Вопрос:

Имеются некоторые данные о переменных  $x_1$ ,  $x_2$  и  $y$  (см. таблицу). Определить стандартизированные коэффициенты регрессии и коэффициенты эластичности. Проверить гипотезу о значимости / незначимости коэффициентов уравнения регрессии для уровня значимости 0,05. Данные можно округлить до сотых. Ответ должен содержать развёрнутое решение и пояснение.

$y$	2	3	4	5	16
$x_1$	-2	-1	0	3	2
$x_2$	3	5	3	-2	-2

### Ответ:

Среднее значение для  $y = 6$

СКО  $y = 5,1$

Среднее значение для  $x_1 = 0,4$

СКО  $x_1 = 1,85$

Среднее значение для  $x_2 = 1,4$

СКО  $x_2 = 2,87$

Определены параметры линейной регрессии

$b_2$	-1,4824191	-0,4233474	$b_1$
$s_{b_2}$	2,0016748	3,09797498	$s_{b_1}$

Стандартизированные коэффициенты регрессии:

$b_1' = -0,154$  – показывает, что при увеличении первой независимой переменной на одно стандартное отклонение зависимая переменная уменьшается в среднем на 0,154 стандартного отклонения, при условии фиксированного значения второй переменной.

$b_2' = -0,835$  – означает, что при росте второй независимой переменной на одно стандартное отклонение зависимая переменная снижается на 0,835 стандартного отклонения при фиксированном значении первой переменной.

Вторая переменная ( $b_2' = -0,835$ ) оказывает значительно более сильное отрицательное влияние на зависимую переменную, чем первая ( $b_1' = -0,154$ ). При этом оба фактора снижают значение зависимой переменной, но второй фактор является доминирующим.

Коэффициенты эластичности:

$E_1 = -0,028$  – означает, что при увеличении первой независимой переменной на 1% зависимая переменная в среднем уменьшается на 0,028% (очень слабое влияние).

$E_2 = -0,346$  – показывает, что рост второй независимой переменной на 1% приводит к снижению зависимой переменной на 0,346% (умеренно сильное влияние).

Вторая переменная ( $E_2 = -0,346$ ) оказывает более существенное влияние в процентном выражении, чем первая ( $E_1 = -0,028$ ). Однако оба коэффициента эластичности по модулю меньше 1, что говорит о неэластичности зависимой переменной по отношению к данным факторам (изменения факторов слабо влияют на результат).

Основной фактор, влияющий на зависимую переменную – вторая независимая переменная ( $b_2' = -0,835$ ,  $E_2 = -0,346$ ), которая имеет значительно большее влияние, чем первая. Характер связи – отрицательный, так как рост обеих переменных приводит к снижению зависимой переменной. Если целью является управление зависимой переменной, следует в первую очередь учитывать влияние второй переменной, так как ее эффект наиболее значим. Первая переменная практически не играет роли в модели.

Проверка гипотезы о незначимости коэффициентов уравнения регрессии для уровня значимости 0,05.

Критическое значение (порог) Т-статистики:  $t = 4.3$

$T_1 = -0,136$

$T_2 = -0,74$

Поскольку  $|T_1| < 4,3$  и  $|T_2| < 4,3$ , ни один из коэффициентов не является статистически значимым на выбранном уровне. Это означает, что нет достаточных оснований утверждать, что переменные  $x_1$  и  $x_2$  действительно влияют на зависимую переменную  $y$  в данной модели.

Ни один из коэффициентов регрессии не является статистически значимым при заданном критическом значении  $t = 4,3$ . Это означает, что в рамках данной модели нет доказательств влияния переменных  $x_1$  и  $x_2$  на  $y$ . Для дальнейшего анализа требуется либо больше данных, либо пересмотр модели (например, включение других факторов или изменение уровня значимости).

### 2.2. Вопрос:

Имеются некоторые данные о переменных  $x_1$ ,  $x_2$  и  $y$  (см. таблицу). Определить стандартизированные коэффициенты регрессии и коэффициенты эластичности. Проверить гипотезу о значимости / незначимости коэффициентов уравнения регрессии для уровня значимости 0,05. Данные можно округлить до сотых. Ответ должен содержать развёрнутое решение и пояснение.

y	5	10	10	7	5	6	6	5
x1	8	11	12	9	8	8	9	9
x2	5	8	8	5	7	8	6	4

### 2.3. Вопрос:

Имеются некоторые данные о переменных  $x_1$ ,  $x_2$ ,  $x_3$ ,  $x_4$  и  $y$  (см. таблицу). Определить стандартизированные коэффициенты регрессии и коэффициенты эластичности. Проверить гипотезу о значимости / незначимости коэффициентов уравнения регрессии для уровня значимости 0,05. Данные можно округлить до сотых. Ответ должен содержать развёрнутое решение и пояснение.

y	1	2	3	4	5	6	7	8
x1	2	2	3	3	2	4	2	2
x2	2	2	1,5	2	3	2	1,5	2
x3	20	12	33	43	53	23	99	34
x4	142	144	151	150	139	169	126	142,9

### 2.4. Вопрос:

Имеются некоторые данные о переменных  $x_1$ ,  $x_2$  и  $y$  (см. таблицу). Определить стандартизированные коэффициенты регрессии и коэффициенты эластичности. Проверить гипотезу о значимости / незначимости коэффициентов уравнения регрессии для уровня значимости 0,05. Данные можно округлить до сотых. Ответ должен содержать развёрнутое решение и пояснение.

y	6,8	6,2	5,9	5,7	5,5	5,3	4,3
x1	4,1	5,9	5,2	6,8	5,8	4,2	5,2
x2	5,1	5,9	7,2	1,8	8,8	7,2	5,2

### Задание 3 (5 баллов)

#### 3.1. Вопрос:

Структура модели ARIMA. Параметры модели ARIMA и процесс их определения. Ответ должен содержать развёрнутое решение и пояснение.

#### Ответ:

Среди статистических моделей для анализа временных рядов наиболее часто используются следующие модели: AR (авторегрессивная), MA (модель скользящего среднего), ARMA (модель авторегрессии — скользящего среднего) и ARIMA (интегрированная модель авторегрессии — скользящего среднего).

AR представляет из себя регрессию или прогноз, основанный на прошлых значениях из того же временного ряда. Из-за своих особенностей значения не меняются молниеносно.

$$y_t = \beta_1 + \rho_1 y_{t-1} + \rho_2 y_{t-2} + \dots + \rho_p y_{t-p} + \varepsilon_t, \quad t = 1, \dots, T$$

В авторегрессионных (AR) моделях текущее значение процесса представляется как линейная комбинация предыдущих его значений и случайной компоненты. В AR(p) – модели авторегрессии порядка p текущий уровень ряда представляется в виде взвешенной суммы p предыдущих наблюдений.

Во многих случаях AR-модель оказывается перегруженной незначимыми коэффициентами, которые надо исключить.

AR-модели не предназначены для описания процессов с тенденцией, однако они хорошо описывают колебания, что весьма важно для отображения развития неустойчивых показателей. Чтобы сделать возможным применение AR-моделей к процессам с тенденцией, на первом этапе формируют стационарный ряд, т. е. исключают тенденцию.

Идентификация AR(p) модели состоит в определении ее порядка p. В сезонной модели авторегрессии порядок выбирается равным периоду сезонности. При отсутствии сезонности начальная оценка порядка параметра p формируется на основе анализа автокорреляционной функции.

MA даёт значения, основанные на ошибке предыдущих прогнозов.

Модель скользящего среднего предполагает, что в ошибках модели в предшествующие периоды сосредоточена информация обо всей предыстории ряда.

Модель скользящего среднего является распространенным подходом для моделирования одномерных временных рядов. Согласно модели, оценка прогнозируемых членов ряда линейно зависит от текущего и прошлых значений, а также некоторого стохастического члена, который отражает вероятностный характер модели.

Модель скользящего среднего представляет частный случай более общей модели авторегрессии скользящего среднего, которая наиболее полно учитывает вероятностные свойства ряда.

Модель скользящего среднего порядка q записывается в виде:

$$X_t = \mu + \varepsilon_t + \Theta_1 \varepsilon_{t-1} + \dots + \Theta_q \varepsilon_{t-q},$$

где  $\mu$  — среднее значение ряда,  $\Theta_1, \dots, \Theta_q$  — параметры модели,  $\varepsilon_t, \varepsilon_{t-1}, \dots$

— шумовые компоненты.

Таким образом, модель скользящего среднего концептуально представляет собой линейную регрессию прогнозируемого значения ряда относительно текущего и ранее наблюдаемых случайных компонентов. Предполагается, что случайные члены взаимно независимы, подчиняются одному и тому же распределению, обычно нормальному, с нулевым математическим ожиданием.

Оценка параметров моделей скользящего среднего сложнее, чем авторегрессионных, поскольку вместо метода наименьших квадратов использует сложные итерационные процедуры.

ARMA — модель, комбинирующая AR и MA. Прогнозирует значения, основываясь и на предыдущих значениях, и на ошибках прогноза.

В общем виде модель ARMA(p,q), где p – порядок авторегрессии, q - порядок скользящего среднего, выглядит следующим образом:

$$y_t = \alpha_1 y_{t-1} + \dots + \alpha_p y_{t-p} + \varepsilon_t + \theta_1 y_{t-1} + \dots + \theta_q y_{t-q} + \varepsilon_t$$

ARIMA дополняет модель ARMA интегрированием, позволяющим нивелировать воздействие неожиданных всплесков в данных интегрированием. Используется чаще всего, так как позволяет делать прогнозы на основе нестационарных временных рядов (рядов, в которых нет четкой детерминированности поведения).

Если процесс оказывается нестационарным и для приведения его к стационарному виду потребовалось взять несколько разностей, то модель становится ARIMA(p,d,q), где d – порядок разности.

$$y_t = \beta_1 y_{t-1} + \beta_2 y_{t-2} + \beta_3 y_{t-3} + \dots + \beta_p y_{t-p} + \varepsilon_t + \alpha_1 \varepsilon_{t-1} + \alpha_2 \varepsilon_{t-2} + \dots + \alpha_q \varepsilon_{t-q}$$

Параметры ARIMA(p,d,q)

Параметр d — это порядок дифференцирования временного ряда, при котором ряд становится стационарным. Стационарность определяется по характеру графика АКФ (функции, выясняющей периодичность сигнала): если график сходится к нулевому значению, значит, временной ряд стационарен.

Параметр p представляет собой порядок авторегрессии. Он равен количеству значений автокорреляции, выходящих за доверительный интервал на графике АКФ отдифференцированного ряда.

Параметр q отвечает за члены скользящего среднего. Он выясняется при анализе ЧАКФ отдифференцированного ряда и равен количеству значений, выходящих за доверительный интервал.

Идентификация модели стационарного временного ряда предназначена для предварительного определения типа модели (авторегрессия, скользящее среднее или смешанный процесс) и порядка модели (значения p и q). Однако автокорреляционная функция стационарного временного ряда не позволяет однозначно идентифицировать модель ряда. Это возможно с использованием второй дополнительной функции - частной автокорреляционной функции (ЧАКФ в русской транскрипции или PACF в английской транскрипции). Значения ЧАКФ — это значение m-го коэффициента в представлении временного ряда процессом авторегрессии порядка m.

### 3.2. Вопрос:

В каких ситуациях ARIMA даёт плохие прогнозы? Сравните её с альтернативами. Приведите пример ряда, где ARIMA проигрывает другим методам. Ответ должен содержать развёрнутое решение и пояснение.

### 3.3. Вопрос:

Нелинейные расширения ARIMA. В каких случаях их применение предпочтительнее стандартной ARIMA? Ответ должен содержать развёрнутое решение и пояснение.

### 3.4. Вопрос:

В чём отличие ARIMAX от ARIMA? Как включить внешние регрессоры в модель? Приведите пример уравнения ARIMAX(1,1,1) с одним экзогенным фактором. Какие проблемы могут возникнуть при интерпретации коэффициентов? Ответ должен содержать развёрнутое решение и пояснение.

### Задание 4 (10 баллов)

#### 4.1. Вопрос:

В таблице представлены данные спроса на товар. Сгладить временной ряд методом скользящего среднего, самостоятельно подобрав размер k-окна сглаживания. Рассчитать тренд, сезонность, случайную составляющую и прогноз Y на следующий период. Данные можно округлить до сотых. Ответ должен содержать развёрнутое решение и пояснение.

1	2	3	4	5	6	7	8	9	10
2	4	3	2,5	4,5	3	2,7	4,7	3,5	3,2

#### Ответ:

По данным из таблицы было выявлено, что представлена аддитивная модель временного ряда, которая содержит циклическую компоненту.

Период цикл компоненты = 3.

Первое приближение тренда методом скользящего среднего (u1), равным периоду цикл составляющей.

X	1	2	3	4	5	6	7	8	9	10
Y	2	4	3	2,5	4,5	3	2,7	4,7	3,5	3,2
U1		3	3,17	3,33	3,33	3,4	3,47	3,63	3,8	

Второе приближение цикл составляющей (u2).

X	1	2	3	4	5	6	7	8	9	10
Y	2	4	3	2,5	4,5	3	2,7	4,7	3,5	3,2
U1		3	3,17	3,33	3,33	3,4	3,47	3,63	3,8	
U2			3,08	3,25	3,33	3,37	3,43	3,55	3,72	

Первое приближение циклической компоненты (s1).

X	1	2	3	4	5	6	7	8	9	10
Y	2	4	3	2,5	4,5	3	2,7	4,7	3,5	3,2
U1		3	3,17	3,33	3,33	3,4	3,47	3,63	3,8	
U2			3,08	3,25	3,33	3,37	3,43	3,55	3,72	
s1			-0,08	-0,75	1,17	-0,37	-0,73	1,15		

Второе приближение циклической компоненты (s2).

X	1	2	3	4	5	6	7	8	9	10
Y	2	4	3	2,5	4,5	3	2,7	4,7	3,5	3,2
U1		3	3,17	3,33	3,33	3,4	3,47	3,63	3,8	
U2			3,08	3,25	3,33	3,37	3,43	3,55	3,72	
s1			-0,08	-0,75	1,17	-0,37	-0,73	1,15		
s2			-0,22	-0,74	1,16					

Третье приближение циклической компоненты (s3).

Среднее значение s2 = 0,065

X	1	2	3	4	5	6	7	8	9	10
Y	2	4	3	2,5	4,5	3	2,7	4,7	3,5	3,2
U1		3	3,17	3,33	3,33	3,4	3,47	3,63	3,8	
U2			3,08	3,25	3,33	3,37	3,43	3,55	3,72	
s1			-0,08	-0,75	1,17	-0,37	-0,73	1,15		
s2			-0,22	-0,74	1,16					
s3			-0,29	-0,81	1,09					

Оценка параметров тренда по МНК:  $T=(m^{\wedge}) * x+(b^{\wedge})$

$m^{\wedge}= 0,09$

$b^{\wedge}=2,8$

X	1	2	3	4	5	6	7	8	9	10
Y	2	4	3	2,5	4,5	3	2,7	4,7	3,5	3,2
U1		3	3,17	3,33	3,33	3,4	3,47	3,63	3,8	
U2			3,08	3,25	3,33	3,37	3,43	3,55	3,72	
s1			-0,08	-0,75	1,17	-0,37	-0,73	1,15		
s2			-0,22	-0,74	1,16					

s3			-0,29	-0,81	1,09					
<b>T</b>	<b>2,9</b>	<b>2,99</b>	<b>3,08</b>	<b>3,17</b>	<b>3,26</b>	<b>3,36</b>	<b>3,45</b>	<b>3,54</b>	<b>3,63</b>	<b>3,72</b>

Тренд для прогноза:

$$T_{11} = 3,81$$

$$T_{12} = 3,9$$

$$T_{13} = 3,99$$

Прогноз  $y$  на следующий промежуток времени:

$$y_{11} = 4,9$$

$$y_{12} = 3,61$$

$$y_{13} = 3,18$$

#### 4.2.Вопрос:

В таблице представлены данные спроса на товар. Сгладить временной ряд методом скользящего среднего, самостоятельно подобрав размер  $k$ -окна сглаживания. Рассчитать тренд, сезонность, случайную составляющую и прогноз  $Y$  на следующий период. Данные можно округлить до сотых. Ответ должен содержать развёрнутое решение и пояснение.

1	2	3	4	5	6	7	8
2	3	3,5	3,7	2,5	3,3	3,7	4

#### 4.3.Вопрос:

В таблице представлены данные спроса на товар. Сгладить временной ряд методом скользящего среднего, самостоятельно подобрав размер  $k$ -окна сглаживания. Рассчитать тренд, сезонность, случайную составляющую и прогноз  $Y$  на следующий период. Данные можно округлить до сотых. Ответ должен содержать развёрнутое решение и пояснение.

1	2	3	4	5	6	7	8	9	10	11	12
1,5	0,7	1,2	2,9	1,4	1	1,7	2,5	1,9	1,6	2	2,9

#### 4.4.Вопрос:

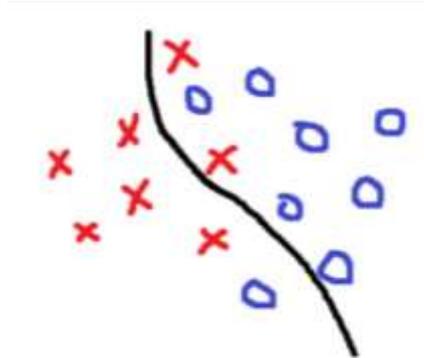
В таблице представлены данные спроса на товар. Сгладить временной ряд методом скользящего среднего, самостоятельно подобрав размер  $k$ -окна сглаживания. Рассчитать тренд, сезонность, случайную составляющую и прогноз  $Y$  на следующий период. Данные можно округлить до сотых. Ответ должен содержать развёрнутое решение и пояснение.

1	2	3	4	5	6	7	8	9	10
2	2,4	3	5	3,2	2,8	2	6	4	1,6

## Задание 5 (15 баллов)

### 5.1. Вопрос:

Дана выборка объектов двух классов (см. рисунок): крестики и кружки. Требуется определить следующие метрики качества классификации относительно кружков класса: Точность (Accuracy), Полнота (Recall), Точность предсказания (Precision), F1-мера (F1-score) и Специфичность (Specificity). Выписать основные определения этих метрик и привести соответствующие формулы. Рассчитать значения указанных метрик, руководствуясь математическими правилами округления до двух знаков после запятой. Написать 2-3 алгоритма машинного обучения для решения задачи классификации. Описать принцип их работы. Ответ должен содержать развёрнутое решение и пояснение.



### Ответ:

Расчеты для синего класса:

- TP (Истинно положительные): 7
- TN (Истинно отрицательные): 5
- FP (Ложно положительные): 2
- FN (Ложно отрицательные): 1

Значения метрик:

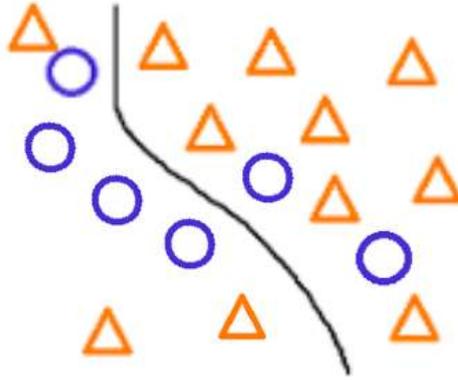
- Точность (Accuracy): 0.80
- Полнота (Recall): 0.88
- Точность предсказания (Precision): 0.78
- F1-мера (F1-score): 0.82
- Специфичность (Specificity): 0.71

Алгоритмы, которые могли бы решить задачу классификации:

- К-ближайших соседей (k-NN).
- Метод опорных векторов (SVM, SVC).
- Деревья решений.
- Случайный лес (Random Forest).
- Градиентный бустинг.
- Логистическая регрессия.
- Нейронные сети.

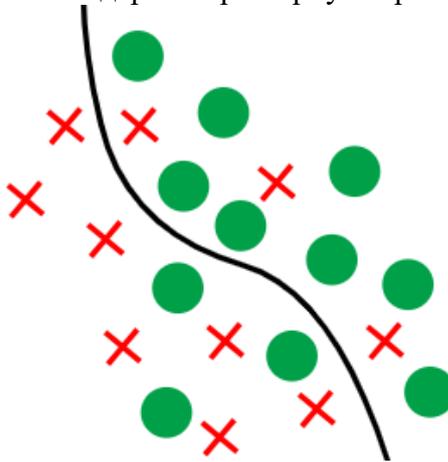
### 5.2. Вопрос:

Дана выборка объектов двух классов: треугольники и кружки. Требуется определить следующие метрики качества классификации относительно кружков: Точность (Accuracy), Полнота (Recall), Точность предсказания (Precision), F1-мера (F1-score) и Специфичность (Specificity). Выписать основные определения этих метрик и привести соответствующие формулы. Рассчитать значения указанных метрик, руководствуясь математическими правилами округления до двух знаков после запятой. Написать 2-3 алгоритма машинного обучения для решения задачи классификации. Описать принцип их работы.



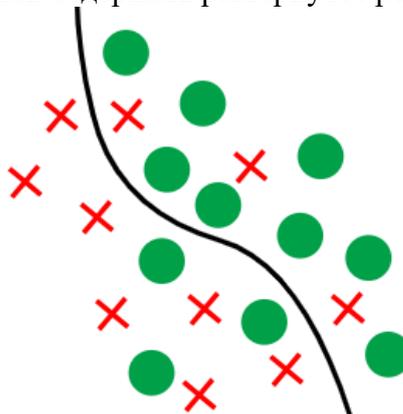
### 5.3. Вопрос:

Дана выборка объектов двух классов: крестики и кружки. Требуется определить следующие метрики качества классификации относительно кружков класса: Точность (Accuracy), Полнота (Recall), Точность предсказания (Precision), F1-мера (F1-score) и Специфичность (Specificity). Выписать основные определения этих метрик и привести соответствующие формулы. Рассчитать значения указанных метрик, руководствуясь математическими правилами округления до двух знаков после запятой. Написать 2-3 алгоритма машинного обучения для решения задачи классификации. Описать принцип их работы. Ответ должен содержать развёрнутое решение и пояснение.



### 5.4. Вопрос:

Дана выборка объектов двух классов: крестики и кружки. Требуется определить следующие метрики качества классификации относительно крестиков класса: Точность (Accuracy), Полнота (Recall), Точность предсказания (Precision), F1-мера (F1-score) и Специфичность (Specificity). Выписать основные определения этих метрик и привести соответствующие формулы. Рассчитать значения указанных метрик, руководствуясь математическими правилами округления до двух знаков после запятой. Написать 2-3 алгоритма машинного обучения для решения задачи классификации. Описать принцип их работы. Ответ должен содержать развёрнутое решение и пояснение.



# Банк заданий по специальной части вступительного испытания в магистратуру

## Задание 6 (5 баллов)

### 6.1. Вопрос:

Понятие модели и моделирования.

### Ответ:

Модель – это упрощённое представление реального объекта, системы или процесса, созданное для изучения, анализа и прогнозирования его свойств и поведения. Модель отражает наиболее существенные характеристики оригинала, игнорируя второстепенные детали.

Ключевые свойства модели:

- Адекватность – модель должна достаточно точно описывать исследуемый объект.
- Упрощение – модель исключает сложные и незначимые детали.
- Целесообразность – модель создаётся для конкретной задачи (прогнозирование, оптимизация, обучение и т. д.).

Типы моделей:

- Материальные (физические) – уменьшенные копии объектов (макеты зданий, модели самолётов в аэродинамической трубе).
- Абстрактные (формальные) – математические, логические, компьютерные модели (уравнения, алгоритмы, графы).
- Вербальные – описательные модели на естественном языке (теории, гипотезы).

Моделирование – это замещение исходного объекта-оригинала другим объектом с целью получения информации о свойствах объекта-оригинала.

Этапы моделирования:

1. Постановка задачи – определение цели моделирования (например, прогнозирование спроса на товар).
2. Выбор типа модели – зависит от природы объекта и целей (математическая, имитационная, графическая).
3. Построение модели – формализация объекта (например, составление уравнений или создание 3D-визуализации).
4. Верификация и валидация – проверка корректности модели (верификация – соответствие формальным требованиям, валидация – соответствие реальному объекту).
5. Использование модели – проведение экспериментов, прогнозирование, оптимизация.
6. Интерпретация результатов – анализ выводов и принятие решений.

Виды моделирования:

- Аналитическое – решение уравнений, описывающих систему (например, расчёт траектории полёта).
- Имитационное – воспроизведение поведения системы во времени (например, экономические симуляторы).
- Статистическое – использование данных для построения вероятностных моделей.
- Компьютерное – применение программных средств (CAD, ANSYS, MATLAB).

Моделирование позволяет:

- Исследовать сложные системы без дорогостоящих экспериментов.
- Прогнозировать последствия решений (например, климатические модели).
- Оптимизировать процессы (логистика, производство).
- Обучать специалистов (тренажёры, симуляторы).

Примеры моделей:

- Математическая модель эпидемии для прогнозирования распространения болезни.
- 3D-модель здания в архитектуре.
- Нейросетевая модель для распознавания изображений.

**6.2.Вопрос:**

Понятие случайной величины, моделирование случайных событий.

**6.3.Вопрос:**

Системы массового обслуживания (СМО). Классификация СМО.

**6.4.Вопрос:**

Основные этапы процесса имитации. Разработка модели и оценка её адекватности.

## Задание 7 (5 баллов)

### 7.1. Вопрос:

Организационные структуры управления проектом.

### Ответ:

Организационная структура проекта определяет распределение ролей, ответственности и коммуникационных потоков между участниками.

Основные типы структур:

#### 1. Функциональная структура

Участники проекта подчиняются своим функциональным руководителям (отделам: разработка, маркетинг, финансы и т. д.). Управление проектом распределено между отделами.

Пример: В IT-компании разработчики подчиняются техническому директору, тестировщики – QA-лиду, а менеджер проекта координирует их работу без прямой власти.

Плюсы:

- Высокая экспертиза в каждой области.
- Минимизация дублирования функций.
- Чёткая карьера для сотрудников.

Минусы:

- Медленное принятие решений (согласование между отделами).
- Слабый контроль менеджера проекта.
- Риск конфликтов приоритетов (проект vs. рутинные задачи).

#### 2. Проектная структура

Создаётся временная команда, которая подчиняется исключительно руководителю проекта. После завершения проекта участники расформируются.

Пример: Строительство моста: инженеры, архитекторы и рабочие работают только на этот проект под управлением РМ.

Плюсы:

- Быстрые решения (единая вертикаль власти).
- Чёткая ответственность за результат.
- Гибкость в распределении ресурсов.

Минусы:

- Неэффективное использование ресурсов между проектами.
- Проблемы с мотивацией (временная занятость).
- Потеря экспертизы после закрытия проекта.

#### 3. Матричная структура

Гибрид функциональной и проектной структур. Участники подчиняются и функциональному руководителю, и менеджеру проекта.

Подтипы матричной структуры:

- Слабая матрица: Власть у функциональных менеджеров, РМ – координатор.
- Сбалансированная матрица: Власть разделена между РМ и функциональными руководителями.
- Сильная матрица: РМ имеет приоритет, функциональные руководители играют вспомогательную роль.

Пример (сбалансированная матрица): Разработка ПО – программисты подчиняются техлиду (функционально) и РМ (по задачам проекта).

Плюсы:

- Баланс между экспертизой и фокусом на проект.
- Гибкость в использовании ресурсов.

Минусы:

- Конфликты из-за двойного подчинения.
- Сложность управления приоритетами.

Можно выделить и другие организационные структуры управления:

- Agile-структура (скрам, канбан) – самоорганизующиеся кросс-функциональные команды без жёсткой иерархии.
- Гибридная структура, которая комбинирует элементы разных структур для адаптации под специфику проекта.

Выбор структуры зависит от:

- Сложности проекта (например, для инноваций подойдёт Agile).
- Корпоративной культуры (матричная структура требует зрелого менеджмента).
- Внешних условий (при жёстких сроках лучше подходит проектная структура).

#### **7.2. Вопрос:**

Эволюционные стратегии разработки.

#### **7.3. Вопрос:**

Прототипирование в разработке ПО.

#### **7.4. Вопрос:**

Спиральная модель Бозма.

## Задание 8 (10 баллов)

### 8.1. Вопрос:

Принципы SOLID.

### Ответ:

SOLID — это набор пяти принципов проектирования.

#### 1. Принцип единственной ответственности (Single Responsibility Principle, SRP):

Класс должен иметь только одну причину для изменения, то есть выполнять лишь одну задачу.

Пример нарушения:

```
public class Order
{
    public void CalculateTotal() { /* Логика расчёта суммы заказа */ }
    public void PrintInvoice() { /* Логика печати счёта */ }
}
```

Правильный подход – разделить класс Order на два класса.

#### 2. Принцип открытости/закрытости (Open/Closed Principle, OCP):

Классы должны быть открыты для расширения, но закрыты для модификации.

Пример нарушения:

```
public double CalculateArea(object shape)
{
    if (shape is Rectangle) { /* Расчёт площади прямоугольника */ }
    else if (shape is Circle) { /* Расчёт площади круга */ }
    // При добавлении новой фигуры придётся изменять этот метод
}
```

Правильный подход – использовать интерфейс IShape с методом CalculateArea и наследование.

#### 3. Принцип подстановки Барбары Лисков (Liskov Substitution Principle, LSP):

Подклассы должны заменять свои базовые классы, не ломая логику программы.

Пример нарушения:

```
public class Bird
{
    public virtual void Fly() { /* ... */ }
}
public class Penguin : Bird
{
    public override void Fly() => throw new NotSupportedException("Пингвины не летают!");
}
```

Правильный подход – исключить наследование Penguin от Bird. Ввести более специализированные интерфейсы.

#### 4. Принцип разделения интерфейсов (Interface Segregation Principle, ISP):

Программные сущности не должны зависеть от методов, которые они не используют. Интерфейсы не должны включать в себя слишком много методов.

Пример нарушения:

```
public interface IWorker
{
    void Work();
    void Eat();
}

public class Robot : IWorker
{
    public void Work() { /* ... */ }
    public void Eat() => throw new NotSupportedException("Роботы не едят!");
}
```

Правильный подход – разделить интерфейс IWorker. Ввести более специализированные интерфейсы.

5. Принцип инверсии зависимостей (Dependency Inversion Principle, DIP):

Модули верхнего уровня не должны зависеть от модулей нижнего уровня. Оба должны зависеть от абстракций.

Пример нарушения:

```
public class OrderService
{
    private readonly SqlDatabase _database; // Жёсткая зависимость от конкретной БД
}
```

Правильный подход – реализовать интерфейс IDatabase и использовать его наследника SqlDatabase.

Применение принципов SOLID позволяет сделать код более гибким (легко добавлять новую функциональность), поддерживаемым (меньше связанности, проще тестировать), масштабируемым (удобно расширять без переписывания).

## 8.2. Вопрос:

Задачи анализа данных Data Mining.

### Ответ:

Data Mining (интеллектуальный анализ данных) – это процесс обнаружения скрытых закономерностей, аномалий и полезных знаний в больших массивах данных с использованием методов машинного обучения, статистики и алгоритмов.

Основные задачи Data Mining можно разделить на описательные (анализ текущих данных) и предсказательные (прогнозирование будущих событий).

Описательные задачи:

1. Кластеризация (Clustering) – разделение данных на группы (кластеры) по схожести без заранее известных меток.

Применение:

- Сегментация клиентов (например, для персонализации рекламы).
- Группировка документов по темам.

2. Ассоциативные правила (Association Rules) – поиск взаимосвязей между событиями или объектами в транзакционных данных.

Применение:

Анализ покупок в супермаркете (правило "Если купили угли для шашлыка, то часто берут смесь для розжига").

Рекомендательные системы (Amazon, Netflix).

3. Анализ аномалий (Anomaly Detection) – обнаружение нестандартных данных (выбросов).

Применение:

Выявление мошеннических транзакций.

Обнаружение сбоев в промышленном оборудовании.

Предсказательные задачи

1. Классификация (Classification) – отнесение объектов к заранее определённым классам.

Применение:

Определение спама в email.

Диагностика заболеваний по симптомам.

2. Регрессия (Regression) – прогнозирование числовых значений на основе исторических данных.

Применение:

Предсказание цен на недвижимость.

Прогнозирование спроса на товары.

3. Прогнозирование временных рядов (Time Series Forecasting) – предсказание будущих значений на основе временных данных.

Применение:

Прогноз курса акций.

Планирование нагрузки на серверы.

Прочие задачи:

1. Снижение размерности (Dimensionality Reduction) – упрощение данных без потери ключевой информации.

Применение:

Визуализация многомерных данных.

Ускорение обучения моделей.

2. Анализ текстов (Text Mining) – извлечение смысла из неструктурированных текстов.

Применение:

Анализ отзывов клиентов.

Классификация новостей.

Data Mining решает широкий спектр задач: от поиска скрытых закономерностей до точных прогнозов. Методы Data Mining применяются в финансах, медицине, ритейле, IT и других областях.

### **8.3.Вопрос:**

Машинный язык. Ассемблер. Языки программирования высокого уровня.

### **8.4.Вопрос:**

Структурное программирование. Принципы структурного подхода к разработке ПО.

### **8.5.Вопрос:**

Задача кластеризации. Применяемые методы.

### **8.6.Вопрос:**

Деревья решений. Процесс конструирования.



## Задание 9 (10 баллов)

### 9.1. Вопрос:

Понятие БД. Основные свойства современных СУБД.

### Ответ:

База данных (БД) – это организованная, структурированная совокупность данных, хранящихся в электронном виде и управляемых с помощью специализированного программного обеспечения – системы управления базами данных (СУБД). СУБД – это совокупность языковых и программных средств, предназначенных для создания, ведения и совместного использования БД многими пользователями.

Ключевые характеристики БД (обеспечиваются СУБД):

- Структурированность – данные организованы по определённым правилам (таблицы, документы, графы и т. д.).
- Постоянность – данные сохраняются после завершения работы программы.
- Доступность – возможность извлечения и модификации данных через интерфейсы (SQL, API и др.).
- Целостность – данные соответствуют заданным правилам и ограничениям.

Современные системы управления базами данных (СУБД) обладают следующими ключевыми свойствами:

- Надёжность и отказоустойчивость

Транзакции (ACID) – гарантия атомарности, согласованности, изолированности и долговечности операций.

Репликация – автоматическое копирование данных на несколько серверов для защиты от сбоев (например, в PostgreSQL можно настроить репликацию master-slave).

Резервное копирование – возможность восстановления данных после аварии.

- Масштабируемость

Горизонтальное масштабирование – распределение нагрузки между несколькими серверами (например, в Postgres).

Вертикальное масштабирование – увеличение мощности одного сервера (например, добавление RAM в SQL-сервер).

- Производительность

Индексация – ускорение поиска данных (B-деревья, хеш-индексы).

Кэширование – хранение часто используемых данных в памяти (например, Redis).

Оптимизация запросов – автоматический выбор эффективного плана выполнения (например, в Postgres).

- Безопасность

Аутентификация и авторизация – контроль доступа пользователей (роли, права).

Шифрование данных – защита информации при хранении и передаче.

Аудит – журналирование действий пользователей.

- Поддержка различных моделей данных

Реляционные (SQL) – таблицы с жёсткой схемой (PostgreSQL, MySQL).

NoSQL – документы (MongoDB), ключ-значение (Redis), графы (Neo4j).

Гибридные — поддержка SQL и NoSQL (например, PostgreSQL с JSONB).

- Распределённость

Шардирование – разделение данных между серверами для балансировки нагрузки.

Поддержка интеграций с облачными решениями.

- **Расширяемость**

Поддержка пользовательских функций – возможность добавлять свои процедуры на различных языках (например, T-SQL, PL/pgSQL).

Плагины и расширения — дополнительные модули для специфичных задач (например, в PostgreSQL).

**9.2.Вопрос:**

Уровни абстракции БД. Реляционная СУБД.

**9.3.Вопрос:**

Реляционные СУБД. Понятия домен, кортеж, отношения.

**9.4.Вопрос:**

Реляционные и нереляционные БД.

## Задание 10 (10 баллов)

### 10.1. Вопрос:

Язык UML. Диаграмма состояний. Пример для предметной области «Магазин».

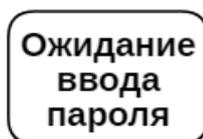
### Ответ:

UML – это стандартизированный графический язык моделирования, используемый для визуализации, спецификации, конструирования и документирования систем.

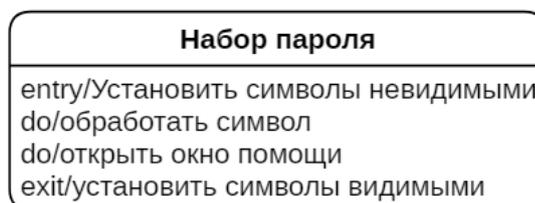
Язык включает более десяти видов диаграмм, одной из которых является диаграмма состояний (диаграмма конечного автомата). Диаграмма состояний (state diagram) показывает автомат, сосредотачивая внимание на потоке управления от одного состояния к другому. Изображается в виде графа с вершинами и дугами (ребрами). Автомат (state machine) – это описание последовательности состояний, через которые проходит объект на протяжении жизненного цикла, реагируя на события, а также описание реакции на эти события.

Основные элементы диаграммы состояний:

Состояния – стабильные условия, в которых может находиться объект. Это ситуация в жизненном цикле объекта, на протяжении которой он удовлетворяет некоторому условию, выполняет некоторую деятельность или ожидает некоторого события. Состояние (простой) изображается прямоугольником с закругленными углами с именем состояния, изображенным внутри этого прямоугольника:



Возможен вариант изображения с разделением на несколько секций. Секциями простого состояния являются: секция имени, секция внутренней деятельности, секция внутренних переходов. Пример изображения:



Деятельность (activity) специфицирует работу, происходящую внутри автомата. Действия – операции, выполняемые при переходе или в состоянии. Это примитивное выполняемое вычисление, приводящее к смене состояния модели или возврату значения.

Начальное псевдосостояние представляет собой частный случай состояния, которое не содержит никаких внутренних действий. Изображается как чёрный круг:



Финальное состояние – специальный вид состояния, предназначенного для моделирования завершения конечного автомата или региона, в котором оно содержится. Изображается как чёрный круг в белой рамке:

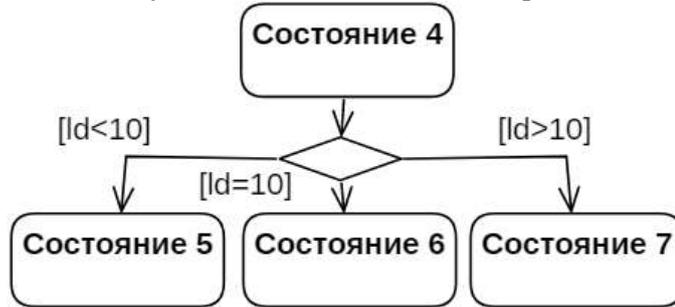


События – триггеры, вызывающие переходы. Это спецификация существенного факта, который происходит во времени и пространстве. В контексте автомата событие – это воздействие, которое вызывает переход между состояниями.

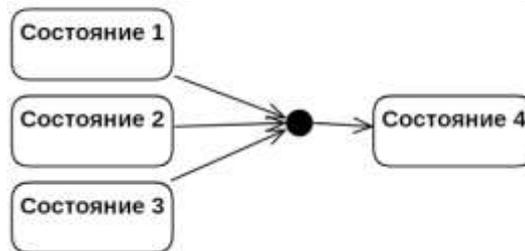
Переходы – изменения состояний под действием событий. Это связь между двумя состояниями, показывающая, что объект, находящийся в первом состоянии, должен выполнить некоторые действия и перейти во второе, как только произойдет определенное событие и будут выполнены определенные условия. Переходы отображаются сплошной линией со стрелкой:



Псевдосостояние выбора предназначено для моделирования нескольких альтернативных ветвей при реализации поведения конечного автомата. Псевдосостояние выбора также называют вершиной или узлом выбора, или просто – выбором. Узел выбора изображается в форме символа ромба. Каждый выходящий из узла выбора переход должен иметь обязательное сторожевое условие, которое представляет собой логическое условие, записанное в квадратных скобках. Пример:

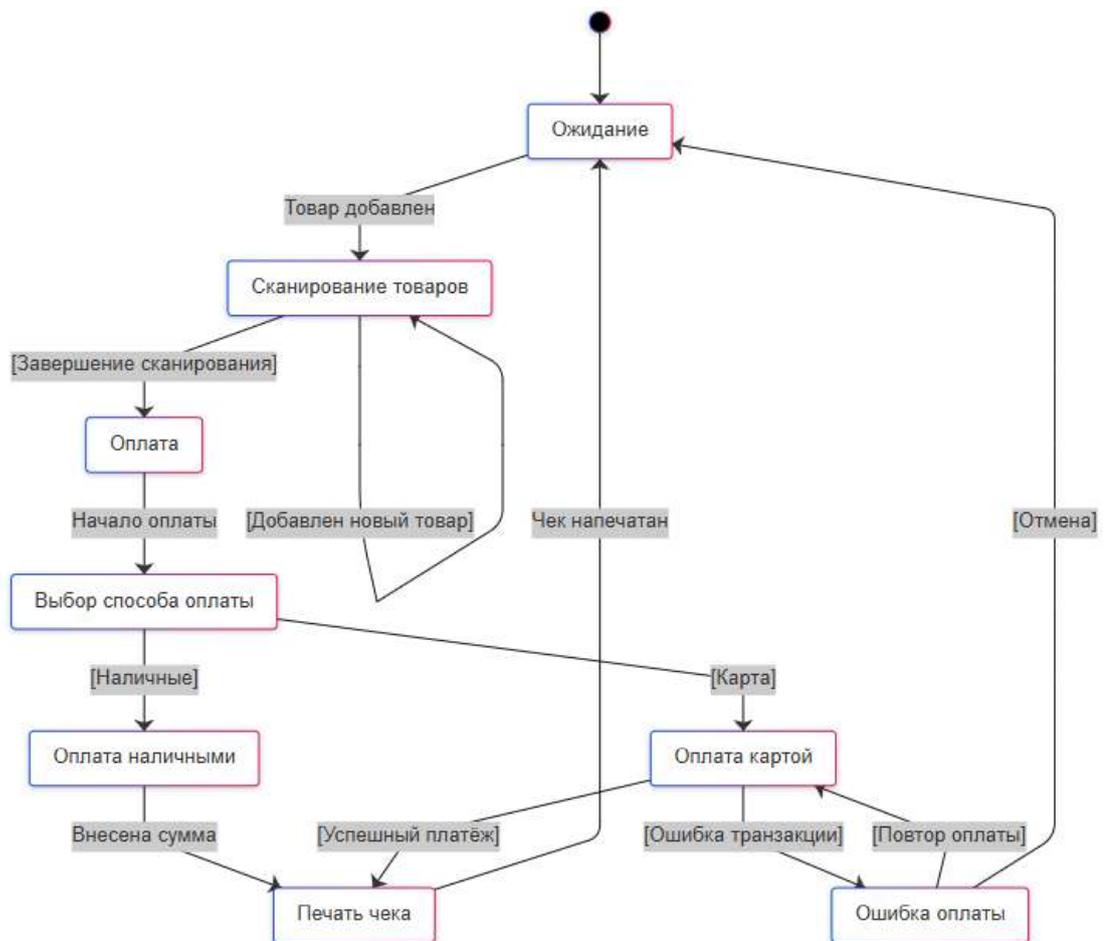


Псевдосостояние соединения является вершиной, которая используется для соединения вместе нескольких переходов. Графически изображается следующим образом:



UML позволяет использовать и более сложные нотации на диаграмме состояний. Например, композитные состояния, в которых присутствуют ортогональные (параллельные) регионы; вершины слияния для описания переходов связанных с ортогональными регионами; точки входа и выхода «из» и «в» композитное состояние.

Пример диаграммы состояний для предметной области «Магазин». Диаграмма охватывает основной цикл работы кассового аппарата: от сканирования товаров до печати чека.



### 10.2. Вопрос:

Язык UML. Диаграмма классов. Пример для предметной области «Университет».

### 10.3. Вопрос:

Язык UML. Диаграмма последовательности. Пример для предметной области «Библиотека».

### 10.4. Вопрос:

Язык UML. Диаграмма вариантов использования. Пример для предметной области «Аквапарк».

### 10.5. Вопрос:

Язык UML. Диаграмма деятельности. Пример для предметной области «Автовокзал».

## Задание 11 (10 баллов)

### 11.1. Вопрос:

Целостность БД. Виды целостности.

### Ответ:

Целостность базы данных – это свойство БД, означающее корректность, непротиворечивость и точность хранимых данных в соответствии с заданными правилами и ограничениями. Целостность обеспечивается механизмами СУБД и предотвращает: нарушение логических связей между данными, появление некорректных значений, потерю согласованности информации.

Выделяют физическую и логическую целостность БД.

Физическая целостность предполагает сохранность данных на уровне носителя. Она обеспечивает защиту от аппаратных сбоев (например, за счёт журналирования транзакций и резервного копирования).

Логическая целостность делится ещё на несколько видов:

1. Целостность сущностей (Entity Integrity) – гарантирует уникальность и обязательность первичного ключа

Правила:

- PRIMARY KEY не может быть NULL.
- Значения PK должны быть уникальны.

Пример нарушения: дубли значений в первичных ключах.

2. Референциальная целостность (Referential Integrity) – контролирует корректность внешних ключей

Правила:

- FOREIGN KEY должен ссылаться на существующий PK
- Обработки изменений, связанных с внешними ключами, одним из вариантов:
  - CASCADE – каскадное обновление/удаление.
  - SET NULL/SET DEFAULT – установка для внешнего ключа значения по умолчанию.
  - RESTRICT – запрет изменений.

Пример нарушения: внешний ключ ссылается на не существующую строку.

3. Доменная целостность (Domain Integrity) – контроль допустимых значений атрибутов.

Правила (механизмы):

- Типы данных (INTEGER, VARCHAR и т.д.).
- Ограничения (NOT NULL, CHECK).
- Значения по умолчанию.

Пример нарушения: отрицательная цена товара.

4. Пользовательская целостность (User-defined Integrity) – бизнес-правила, специфичные для предметной области.

Правила (механизмы)

- Реализация специфичных ограничений через триггеры, хранимые процедуры и аналогичные возможности СУБД.

Пример нарушения: дата доставки товара раньше даты оформления.

5. Отдельным вопросом, связанным с логической целостностью, является операционная целостность (Transactional Integrity). В отличие от статических видов целостности, рассмотренных выше, она решает динамические проблемы, возникающие в многопользовательской среде, и гарантирует согласованность данных при параллельных операциях.

Операционная целостность базируется на четырех ключевых свойствах (аббревиатура ACID):

- Атомарность (Atomicity) – транзакция выполняется как единое целое ("всё или ничего").

- Согласованность (Consistency) – данные переходят из одного корректного состояния в другое.
- Изолированность (Isolation) – параллельные транзакции не влияют друг на друга.
- Долговечность (Durability) – результаты завершённых транзакций сохраняются даже при сбоях.

Основные механизмы СУБД для обеспечения операционной целостности: различные уровни изоляции транзакций, журналирование изменений, блокировки различных видов объектов БД, поддержка версий данных.

#### **11.2. Вопрос:**

Идентифицирующие и не идентифицирующие связи. Правила ограничения целостности.

#### **11.3. Вопрос:**

Основные концепции реляционных СУБД.

#### **11.4. Вопрос:**

Основные методы управления параллельным доступом в СУБД.

## Задание 12 (10 баллов)

### 12.1. Вопрос:

Разработайте модуль программного приложения, позволяющий передавать данные для авторизации пользователя на сервер. Фрагмент схемы БД предложите самостоятельно. Используйте любую известную Вам технологию разработки ПО.

### Ответ:

Для разработки используем следующие технологии: платформа ASP.NET Core, ORM библиотека Entity Framework Core, СУБД PostgreSQL.

Таблица Users для хранения данных пользователей:

```
CREATE TABLE Users (  
    Id SERIAL PRIMARY KEY,  
    Username VARCHAR(50) UNIQUE NOT NULL,  
    Email VARCHAR(100) UNIQUE NOT NULL,  
    PasswordHash VARCHAR(255) NOT NULL,  
    Salt VARCHAR(100) NOT NULL  
);
```

Модель пользователя в Entity Framework:

```
public class User  
{  
    public int Id { get; set; }  
    public string Username { get; set; }  
    public string PasswordHash { get; set; }  
    public string Salt { get; set; }  
}
```

Контекст Entity Framework для взаимодействия с БД

```
public class AppDbContext : DbContext  
{  
    ...  
    public DbSet<User> Users { get; set; }  
}
```

// Сервис авторизации

```
public class AuthService  
{  
    private readonly AppDbContext _db;  
    private readonly IConfiguration _config;  
  
    public AuthService(AppDbContext db, IConfiguration config)  
    {  
        _db = db;  
        _config = config;  
    }  
  
    // Попытка входа в систему  
    public async Task<string?> Login(string username, string password)  
    {  
        var user = await _db.Users.FirstOrDefaultAsync(u => u.Username == username);  
        if (user == null)  
            return null;  
  
        var inputHash = HashPassword(password, user.Salt);  
        if (inputHash != user.PasswordHash)
```

```

        return null;
        return GenerateJwtToken(user);
    }

    // Методы HashPassword и GenerateJwtToken опущены для упрощения
}

```

Контроллер для обработки запроса пользователя.

```

public class AuthController : ControllerBase
{
    private readonly AuthService _authService;

    public AuthController(AuthService authService)
    {
        _authService = authService;
    }

    [HttpGet]
    public IActionResult Login()
    {
        return View();
    }

    [HttpPost]
    public async Task<IActionResult> Login(string username, string password)
    {
        var token = await _authService.Login(username, password);
        if(token == null)
            return Unauthorized();

        Response.Cookies.Append("jwt", token);
        // Перенаправляем на главную страницу приложения
        return Redirect("/");
    }
}

```

Форма входа:

```

<form method="post" action="/Auth/Login">
  <div>
    <label>Логин:</label>
    <input type="text" name="username" required />
  </div>
  <div>
    <label>Пароль:</label>
    <input type="password" name="password" required />
  </div>
  <button type="submit">Войти</button>
</form>

```

Дополнительно требуется описать параметры запуска веб-приложения, добавить конфигурацию доступа к БД.

Краткое описание работы:

Пользователь заходит на /Auth/Login, заполняет HTML-форму с логином и паролем.

После отправки формы на сервере определяется пользователь; хешируется переданный пароль с солью, которая была ранее сохранена в БД; полученный хеш проверяется на соответствие

хешу из БД. Если хеши совпадают, то генерируется JWT-токен и сохраняется в куки. Если нет – возвращается ошибка.

### **12.2. Вопрос:**

Разработайте модуль программного приложения, позволяющий передавать данные на сервер для добавления товара в корзину в интернет-магазине. Фрагмент схемы БД предложите самостоятельно. Используйте любую известную Вам технологию разработки ПО.

### **12.3. Вопрос:**

Разработайте модуль программного приложения, позволяющий передавать данные на сервер для добавления отзыва о товаре интернет-магазине. Фрагмент схемы БД предложите самостоятельно. Используйте любую известную Вам технологию разработки ПО.

### **12.4. Вопрос:**

Разработайте модуль программного приложения, позволяющий передавать данные на сервер для добавления промокода к заказу в интернет-магазине. Фрагмент схемы БД предложите самостоятельно. Используйте любую известную Вам технологию разработки ПО.