

# Вопросы для подготовки к промежуточной аттестации по курсу "Обработка и анализ данных"

---

## В форме тестирования

### Тестовые вопросы по MaxAbsScaler, MinMaxScaler, StandardScaler, RobustScaler

---

1. Какой метод масштабирования НЕ изменяет форму распределения данных?
2. Какой метод масштабирования наиболее устойчив к выбросам?
3. Какой метод масштабирования приводит данные к нулевому среднему и единичной дисперсии?
4. Какой метод масштабирования использует медиану и межквартильный размах (IQR)?
5. Какой метод масштабирования приводит данные к диапазону  $[-1, 1]$ ?
6. Какой метод масштабирования лучше всего подходит для разреженных данных?
7. Если данные имеют длинный хвост и выбросы, какой метод масштабирования предпочтительнее?
8. Какой метод масштабирования использует минимум и максимум признака?

### Тестовые вопросы по StandardScaler

---

1. StandardScaler преобразует данные так, чтобы:
2. StandardScaler чувствителен к выбросам?
3. Какой атрибут объекта StandardScaler содержит средние значения, вычисленные для каждого признака?
4. В каком случае рекомендуется использовать StandardScaler:
5. Применение StandardScaler эквивалентно (как изменяются данные):
6. Что изменяет StandardScaler в исходных данных:
7. Если данные уже имеют нулевое среднее и единичную дисперсию, применение StandardScaler приводит к:

### Тестовые вопросы по MaxAbsScaler

---

1. MaxAbsScaler масштабирует данные выполняя:

2. После применения `MaxAbsScaler` все значения признаков будут находиться в диапазоне:
3. `MaxAbsScaler` хорошо подходит для данных:
4. В результате `MaxAbsScaler` сохраняет:
5. Если в данных есть только положительные значения, `MaxAbsScaler` будет работать аналогично:
6. Что `MaxAbsScaler` изменяет в исходных данных:
7. Если максимальное абсолютное значение признака равно 1, то `MaxAbsScaler`:

## Тестовые вопросы по `MinMaxScaler`

---

1. `MinMaxScaler` масштабирует данные в диапазон:
2. Что `MinMaxScaler` использует для масштабирования:
3. `MinMaxScaler` чувствителен к выбросам?
4. Какой атрибут объекта `MinMaxScaler` содержит минимальные значения, вычисленные для каждого признака?
5. Что `MinMaxScaler` сохраняет в исходных данных:
6. Какой параметр `MinMaxScaler` позволяет задать нужный диапазон масштабирования?
7. Если данные уже находятся в диапазоне `[0, 1]`, применение `MinMaxScaler`:

## Тестовые вопросы по `RobustScaler`

---

1. Что `RobustScaler` использует для масштабирования:
2. `RobustScaler` менее чувствителен к выбросам, чем:
3. Какой параметр `RobustScaler` контролирует диапазон квантилей для вычисления IQR?
4. `with_centering=False` в `RobustScaler` означает:
5. `RobustScaler` рекомендуется использовать, когда:
6. После применения `RobustScaler` с параметрами по умолчанию, медиана каждого признака будет:
7. IQR (Interquartile Range) вычисляется как:

## Тестовые вопросы по `numpy.mean`, `numpy.std`

---

1. Функция `np.mean()` вычисляет:
2. Функция `np.std()` вычисляет:
3. Для вычисления среднего значения элементов вдоль строк двумерного массива `arr` следует использовать:
4. Для вычисления стандартного отклонения элементов вдоль столбцов двумерного массива `arr` следует использовать:
5. Параметр `ddof` в функции `np.std()` определяет:

6. Если в массиве есть значения `NaN`, `np.mean()` вернет:
7. Чтобы проигнорировать `NaN` при вычислении среднего значения, нужно использовать:
8. Чтобы вычислить взвешенное среднее значение, следует использовать:

## Тестовые вопросы по QuantileTransformer

---

1. QuantileTransformer преобразует данные так, чтобы:
2. Какой параметр QuantileTransformer определяет количество квантилей?
3. Какой параметр QuantileTransformer позволяет выбрать между равномерным и нормальным распределением преобразованных данных?
4. Если `output_distribution='normal'`, QuantileTransformer преобразует данные так, чтобы они приблизительно соответствовали:
5. QuantileTransformer устойчив к выбросам?
6. Параметр `random_state` в QuantileTransformer используется для:
7. Если в данных есть пропущенные значения (NaN), QuantileTransformer:

## Тестовые вопросы по PowerTransformer

---

1. Основная цель PowerTransformer:
2. PowerTransformer использует следующую технику для преобразования данных:
3. Какой параметр PowerTransformer определяет метод преобразования?
4. Значение `'yeo-johnson'` параметра `method` означает:
5. Параметр `standardize` в PowerTransformer:
6. PowerTransformer подходит для данных:
7. Box-Cox transformation можно применять только к:

## Тестовые вопросы по KBinsDiscretizer

---

1. KBinsDiscretizer используется для:
2. Параметр `n_bins` в KBinsDiscretizer определяет:
3. Какие стратегии разбиения на интервалы поддерживает KBinsDiscretizer (параметр `strategy`)?
4. Стратегия `uniform` в KBinsDiscretizer:
5. Стратегия `quantile` в KBinsDiscretizer:
6. Параметр `encode` в KBinsDiscretizer определяет способ кодирования преобразованных признаков:
7. `encode='ordinal'` в KBinsDiscretizer преобразует данные в:
8. KBinsDiscretizer может обрабатывать пропущенные значения (NaN)?

## Тестовые вопросы по PCA

---

1. PCA используется для:
2. PCA находит:
3. Результатом работы PCA являются:
4. Главные компоненты являются:
5. Параметр `n_components` в PCA определяет:
6. Параметр `svd_solver` в PCA определяет:
7. `explained_variance_ratio_` в PCA показывает:
8. Перед применением PCA рекомендуется:

## Тестовые вопросы по методу главных компонент (PCA)

---

1. Основная цель PCA:
2. PCA ищет:
3. Результатом PCA являются:
4. Главные компоненты:
5. `n_components` в PCA определяет:
6. `explained_variance_ratio_` в PCA показывает:
7. Перед применением PCA рекомендуется:
8. PCA чувствителен к масштабу признаков?
9. Какой метод чаще всего используется для вычисления главных компонент в PCA?
10. Что такое собственные значения в контексте PCA?
11. `components_` в PCA содержит:
12. `mean_` в PCA содержит:
13. PCA может использоваться для:
14. Если `n_components` задан как число от 0 до 1, то это:
15. `whiten=True` в PCA:
16. PCA является:
17. `inverse_transform` в PCA:
18. Какой параметр отвечает за случайность в PCA?
19. `svd_solver='arpack'` рекомендуется для:
20. PCA гарантирует:

## Тестовые вопросы по Kernel PCA

---

1. Какая из следующих функций ядра НЕ является допустимой для Kernel PCA?
2. Что из перечисленного является целью Kernel PCA?

3. Какое из следующих утверждений о Kernel PCA является верным?
4. Какая из следующих матриц используется для вычисления собственных значений и собственных векторов в Kernel PCA?
5. Что из перечисленного является преимуществом Kernel PCA?

## Тестовые вопросы по Sparse PCA

---

1. Какой параметр в SparsePCA используется для управления степенью разреженности?
2. Какой метод в SparsePCA используется для вычисления компонент?
3. Какой атрибут в SparsePCA содержит коэффициенты разреженных компонент?
4. Какой метод в SparsePCA используется для восстановления исходных данных из спроецированных данных?
5. Что означает "разреженность" в контексте Sparse PCA?
6. Какой из следующих методов регуляризации чаще всего используется в Sparse PCA?
7. Как влияет увеличение параметра регуляризации (например, alpha в Scikit-Learn) на разреженность компонент в Sparse PCA?
8. В каких случаях Sparse PCA может быть предпочтительнее стандартного PCA?
9. Какой из следующих недостатков характерен для Sparse PCA?

## Тестовые вопросы по "Факторный анализ"

---

1. Какова основная цель факторного анализа?
2. Что такое "фактор" в факторном анализе?
3. Какой тип данных подходит для факторного анализа?
4. Что такое "факторная нагрузка"?
5. Какой метод вращения факторов используется для упрощения интерпретации факторов?
6. Что такое "собственное значение" в факторном анализе?
7. Какое правило используется для определения количества факторов, которые следует оставить?
8. Что такое "общность" (communality) в факторном анализе?
9. Какой из следующих методов НЕ является методом извлечения факторов?
10. Факторный анализ является \_\_\_ методом.

## Тестовые вопросы по

`sklearn.decomposition.FactorAnalysis`

---

1. Какой метод используется в `sklearn.decomposition.FactorAnalysis` для оценки параметров модели?

2. Какой параметр в `FactorAnalysis` отвечает за количество факторов, которые необходимо извлечь?
3. Какой параметр в `FactorAnalysis` управляет типом вращения факторов?
4. Какой атрибут обученного объекта `FactorAnalysis` содержит матрицу факторных нагрузок?
5. Какой метод в `FactorAnalysis` используется для трансформации новых данных с использованием обученной модели?
6. Какой параметр в `FactorAnalysis` отвечает за инициализацию факторных нагрузок?
7. Какой параметр в `FactorAnalysis` позволяет контролировать максимальное количество итераций при оценке параметров?
8. Какой атрибут обученного объекта `FactorAnalysis` содержит ковариационную матрицу шума?
9. Какой параметр в `FactorAnalysis` отвечает за допуск сходимости алгоритма?
10. Какой метод в `FactorAnalysis` используется для вычисления логарифмической вероятности данных?

## Тестовые вопросы по "Ассоциативный анализ"

---

1. Что является основной целью ассоциативного анализа?
2. Какой тип данных чаще всего используется в ассоциативном анализе?
3. Что такое "поддержка" (support) правила в ассоциативном анализе?
4. Что такое "достоверность" (confidence) правила в ассоциативном анализе?
5. Что такое "подъем" (lift) правила в ассоциативном анализе?
6. Какой алгоритм часто используется для поиска ассоциативных правил?
7. Что такое "антецедент" в ассоциативном правиле?
8. Что такое "консеквент" в ассоциативном правиле?
9. Какое значение "подъема" (lift) указывает на независимость антецедента и консеквента?
10. Ассоциативный анализ чаще всего применяется в:

## Тестовые вопросы по

`mlxtend.frequent_patterns.apriori`

---

1. Какой тип данных принимает функция `apriori` из библиотеки `mlxtend.frequent_patterns`?
2. Какой параметр в функции `apriori` отвечает за минимальную поддержку?
3. Какой параметр в функции `apriori` позволяет использовать названия столбцов DataFrame вместо индексов?
4. Какой тип данных возвращает функция `apriori`?
5. Какой столбец в возвращаемом DataFrame содержит найденные часто встречающиеся наборы элементов?

6. Какой столбец в возвращаемом DataFrame содержит значение поддержки для каждого часто встречающегося набора?
7. Как можно отфильтровать результаты, полученные с помощью `apriori`, по длине наборов элементов?
8. Для чего используется параметр `low_memory` в функции `apriori`?
9. Какой параметр в функции `apriori` позволяет задать максимальную длину наборов элементов?
10. Что нужно сделать перед использованием функции `apriori`, если данные представлены в виде списка транзакций с текстовыми элементами?

## ~~Тестовые вопросы по "Ассоциативный анализ"~~

---

- ~~1. Что является основной целью ассоциативного анализа?~~
- ~~2. Какой тип данных чаще всего используется в ассоциативном анализе?~~
- ~~3. Что такое "поддержка" (support) правила в ассоциативном анализе?~~
- ~~4. Что такое "достоверность" (confidence) правила в ассоциативном анализе?~~
- ~~5. Что такое "подъем" (lift) правила в ассоциативном анализе?~~
- ~~6. Какой алгоритм часто используется для поиска ассоциативных правил?~~
- ~~7. Что такое "антецедент" в ассоциативном правиле?~~
- ~~8. Что такое "консеквент" в ассоциативном правиле?~~
- ~~9. Какое значение "подъема" (lift) указывает на независимость антецедента и консеквента?~~
- ~~10. Ассоциативный анализ чаще всего применяется в:~~

## Тестовые вопросы по FPGrowth

---

1. Что является основной задачей алгоритма FPGrowth?
2. Какое основное преимущество FPGrowth по сравнению с Apriori?
3. Какая структура данных используется в алгоритме FPGrowth для хранения информации о транзакциях?
4. Что такое "условная база паттернов" (conditional pattern base) в алгоритме FPGrowth?
5. Что такое "условное FP-дерево" (conditional FP-tree) в алгоритме FPGrowth?
6. Как влияет порядок элементов в транзакциях на построение FP-дерева?
7. Какой параметр определяет минимальную поддержку в алгоритме FPGrowth?
8. FPGrowth, как и Apriori, используется на первом этапе поиска:
9. Какой из следующих недостатков может быть у алгоритма FPGrowth?
10. После нахождения часто встречающихся наборов с помощью FPGrowth, что обычно делается для получения ассоциативных правил?

## Тестовые вопросы по `association_rules`

---

1. Для чего используется функция `association_rules` в библиотеке `mlxtend.frequent_patterns`?
2. Какой тип данных принимает функция `association_rules` в качестве основного аргумента?
3. Какой параметр в функции `association_rules` отвечает за метрику, используемую для оценки правил?
4. Какие метрики можно использовать для оценки ассоциативных правил с помощью функции `association_rules`?
5. Какой параметр в функции `association_rules` отвечает за минимальный порог метрики, при котором правило считается интересным?
6. Какой тип данных возвращает функция `association_rules`?
7. Какие столбцы обычно присутствуют в DataFrame, возвращаемом функцией `association_rules`?
8. Что представляют собой столбцы `antecedents` и `consequents` в DataFrame, возвращаемом функцией `association_rules`?
9. Как можно отфильтровать правила, сгенерированные функцией `association_rules`, по значению метрики (например, lift)?
10. Что нужно сделать перед использованием функции `association_rules`, если у вас есть только исходные транзакционные данные?

## Тестовые вопросы по

### `mlxtend.frequent_patterns.fpmax`

---

1. Что является основной задачей алгоритма FPMaх?
2. Что такое "максимальный часто встречающийся набор элементов" в контексте FPMaх?
3. Какой тип данных принимает функция `fpmax` из библиотеки `mlxtend.frequent_patterns`?
4. Какой параметр в функции `fpmax` отвечает за минимальную поддержку?
5. Какой параметр в функции `fpmax` позволяет использовать названия столбцов DataFrame вместо индексов?
6. Какой тип данных возвращает функция `fpmax`?
7. Какой столбец в возвращаемом DataFrame содержит найденные максимальные часто встречающиеся наборы элементов?
8. Какой столбец в возвращаемом DataFrame содержит значение поддержки для каждого максимального часто встречающегося набора?
9. По сравнению с алгоритмом FPGrowth, FPMaх обычно:
10. Что нужно сделать перед использованием функции `fpmax`, если данные представлены в виде списка транзакций с текстовыми элементами?

## Тестовые вопросы по KMeans

---

1. KMeans - это алгоритм:
2. Что является основной целью алгоритма KMeans?



3. Как KMeans определяет сходство между объектами?
4. Что такое центроиды в алгоритме KMeans?
5. Как KMeans выбирает начальное расположение центроидов?
6. Что такое метод локтя (elbow method) в KMeans?
7. Что такое метод силуэта (silhouette method) в KMeans?
8. Какой из следующих недостатков характерен для KMeans?
9. Как можно уменьшить влияние выбросов на KMeans?
10. KMeans гарантированно находит глобальный оптимум?

## Тестовые вопросы по sklearn KMeans

---

1. Какой класс в Scikit-Learn используется для реализации алгоритма KMeans?
2. Какой параметр в `KMeans` отвечает за количество кластеров?
3. Какой параметр в `KMeans` отвечает за метод инициализации центроидов?
4. Какие значения может принимать параметр `init` в `KMeans`?
5. Какой параметр в `KMeans` отвечает за максимальное количество итераций алгоритма?
6. Какой атрибут обученного объекта `KMeans` содержит метки кластеров для каждой точки данных?
7. Какой атрибут обученного объекта `KMeans` содержит координаты центроидов кластеров?
8. Какой метод в `KMeans` используется для предсказания кластера для новых точек данных?
9. Какой метод в `KMeans` используется для вычисления суммы квадратов расстояний от каждой точки до ближайшего центроида?
10. Какой параметр в `KMeans` позволяет задать случайное состояние для воспроизводимости результатов?

## Тестовые вопросы по "Метод Локтя"

---

1. Для чего используется метод локтя?
2. На каком графике основан метод локтя?
3. Как определяется оптимальное количество кластеров с помощью метода локтя?
4. Является ли метод локтя точным методом определения оптимального количества кластеров?
5. Какие недостатки имеет метод локтя?
6. Какие альтернативные методы можно использовать для определения оптимального количества кластеров?
7. Что такое инерция в контексте метода локтя?
8. Как метод локтя связан с принципом уменьшения размерности?
9. Можно ли использовать метод локтя для оценки качества кластеризации, полученной с помощью алгоритмов, отличных от KMeans?
10. Что делать, если на графике метода локтя нет четко выраженной точки "локтя"?

## Тестовые вопросы по MiniBatchKMeans

---

1. MiniBatchKMeans - это модификация алгоритма KMeans, которая:
2. Какое основное отличие MiniBatchKMeans от KMeans?
3. Как MiniBatchKMeans влияет на скорость работы алгоритма?
4. Как MiniBatchKMeans влияет на точность кластеризации?
5. Какой параметр в `MiniBatchKMeans` отвечает за размер мини-батча?
6. Какой параметр в `MiniBatchKMeans` отвечает за количество итераций по всему набору данных?
7. В каких случаях MiniBatchKMeans может быть предпочтительнее KMeans?
8. MiniBatchKMeans гарантированно находит глобальный оптимум?
9. Можно ли использовать метод локтя для определения оптимального количества кластеров в MiniBatchKMeans?
10. Какой атрибут обученного объекта `MiniBatchKMeans` содержит координаты центроидов кластеров?

## Тестовые вопросы по AgglomerativeClustering

---

1. AgglomerativeClustering - это алгоритм:
2. Какой подход использует AgglomerativeClustering для построения кластеров?
3. Что такое дендрограмма в контексте AgglomerativeClustering?
4. Какой параметр в `AgglomerativeClustering` отвечает за количество кластеров?
5. Какой параметр в `AgglomerativeClustering` отвечает за метод связывания (linkage)?
6. Какие методы связывания (linkage) доступны в `AgglomerativeClustering`?
7. Какой метод связывания (linkage) минимизирует дисперсию внутри кластеров?
8. Какой атрибут обученного объекта `AgglomerativeClustering` содержит метки кластеров для каждой точки данных?
9. Можно ли использовать AgglomerativeClustering для кластеризации данных с большим количеством признаков?
10. AgglomerativeClustering гарантированно находит глобальный оптимум?

## Тестовые вопросы по DBSCAN

---

1. DBSCAN - это алгоритм:
2. Какие два основных параметра используются в DBSCAN?
3. Что такое "основная точка" (core point) в DBSCAN?
4. Что такое "границная точка" (border point) в DBSCAN?
5. Что такое "шумовая точка" (noise point) в DBSCAN?
6. Какой параметр в `DBSCAN` отвечает за радиус окрестности?
7. Какой параметр в `DBSCAN` отвечает за минимальное количество точек в окрестности?

8. Какие преимущества имеет DBSCAN по сравнению с KMeans?
9. Какие недостатки имеет DBSCAN?
10. Какой атрибут обученного объекта `DBSCAN` содержит метки кластеров для каждой точки данных?

## Тестовые вопросы по OPTICS

---

1. OPTICS - это алгоритм:
2. Какое основное преимущество OPTICS по сравнению с DBSCAN?
3. Что такое "достижимость" (reachability distance) в OPTICS?
4. Что такое "основное расстояние" (core distance) в OPTICS?
5. Какой параметр в `OPTICS` отвечает за минимальное количество точек в окрестности?
6. Какой параметр в `OPTICS` отвечает за максимальное расстояние для рассмотрения соседей?
7. Что такое "упорядоченный список" (ordering) в OPTICS?
8. Как можно визуализировать результаты OPTICS?
9. Как можно извлечь кластеры из упорядоченного списка OPTICS?
10. OPTICS чувствителен к выбору параметров?

## Тестовые вопросы по GaussianNB

---

1. GaussianNB - это алгоритм:
2. На каком принципе основан GaussianNB?
3. Какое предположение делает GaussianNB о распределении признаков?
4. Какой параметр в `GaussianNB` отвечает за сглаживание (smoothing)?
5. Для чего используется сглаживание в GaussianNB?
6. Какой метод в `GaussianNB` используется для обучения модели?
7. Какой метод в `GaussianNB` используется для предсказания класса для новых данных?
8. Какой атрибут обученного объекта `GaussianNB` содержит средние значения признаков для каждого класса?
9. Какой атрибут обученного объекта `GaussianNB` содержит дисперсии признаков для каждого класса?
10. GaussianNB подходит для работы с категориальными признаками?

## Тестовые вопросы по MultinomialNB

---

1. MultinomialNB - это алгоритм:
2. На каком принципе основан MultinomialNB?
3. Какое предположение делает MultinomialNB о распределении признаков?

4. Какой параметр в `MultinomialNB` отвечает за сглаживание (smoothing)?
5. Для чего используется сглаживание в `MultinomialNB`?
6. Какой метод в `MultinomialNB` используется для обучения модели?
7. Какой метод в `MultinomialNB` используется для предсказания класса для новых данных?
8. Какой атрибут обученного объекта `MultinomialNB` содержит вероятности признаков для каждого класса?
9. `MultinomialNB` подходит для работы с отрицательными значениями признаков?
10. В каких задачах `MultinomialNB` часто показывает хорошие результаты?

## Тестовые вопросы по ComplementNB

---

1. `ComplementNB` - это алгоритм:
2. На каком принципе основан `ComplementNB`?
3. В чем основное отличие `ComplementNB` от `MultinomialNB`?
4. Для каких типов задач `ComplementNB` может быть предпочтительнее `MultinomialNB`?
5. Какой параметр в `ComplementNB` отвечает за сглаживание (smoothing)?
6. Какой метод в `ComplementNB` используется для обучения модели?
7. Какой метод в `ComplementNB` используется для предсказания класса для новых данных?
8. Какой атрибут обученного объекта `ComplementNB` содержит веса признаков для каждого класса?
9. `ComplementNB` подходит для работы с отрицательными значениями признаков?
10. `ComplementNB` может быть использован для решения задач регрессии?

## Тестовые вопросы по BernoulliNB

---

1. `BernoulliNB` - это алгоритм:
2. На каком принципе основан `BernoulliNB`?
3. Какое предположение делает `BernoulliNB` о распределении признаков?
4. Какой параметр в `BernoulliNB` отвечает за сглаживание (smoothing)?
5. Для чего используется сглаживание в `BernoulliNB`?
6. Какой метод в `BernoulliNB` используется для обучения модели?
7. Какой метод в `BernoulliNB` используется для предсказания класса для новых данных?
8. Какой атрибут обученного объекта `BernoulliNB` содержит вероятности появления признаков в каждом классе?
9. `BernoulliNB` подходит для работы с непрерывными признаками?
10. В каких задачах `BernoulliNB` часто показывает хорошие результаты?

## Тестовые вопросы по GaussianNB, MultinomialNB, ComplementNB, BernoulliNB

1. Какой из алгоритмов наивного Байеса подходит для работы с непрерывными признаками, предполагая их нормальное распределение?
2. Вы работаете над задачей классификации текстов, используя представление "мешок слов" (bag-of-words). Какой алгоритм наивного Байеса будет наиболее подходящим?
3. Ваш набор данных для классификации сильно несбалансирован, один класс значительно преобладает над другими. Какой алгоритм наивного Байеса может быть наиболее эффективен в этой ситуации?
4. Ваши признаки представляют собой бинарные значения (0/1), указывающие на наличие или отсутствие определенного атрибута. Какой алгоритм наивного Байеса следует использовать?
5. Какой параметр используется во всех перечисленных алгоритмах наивного Байеса для сглаживания вероятностей и предотвращения проблем с нулевыми вероятностями?
6. Какой из алгоритмов наивного Байеса использует дополнение (complement) вероятностей при расчете вероятности класса, что делает его более устойчивым к шуму в данных?
7. Вы работаете над задачей классификации изображений, где признаки представляют собой значения пикселей. Какой алгоритм наивного Байеса может быть подходящим выбором?
8. Вы разрабатываете систему для фильтрации спама, где признаки представляют собой наличие или отсутствие определенных слов в электронном письме. Какой алгоритм наивного Байеса может быть наиболее эффективен?
9. Вам нужно предсказать вероятность заболевания, основываясь на наличии или отсутствии определенных симптомов у пациента. Какой алгоритм наивного Байеса можно использовать?
10. Все перечисленные алгоритмы наивного Байеса (GaussianNB, MultinomialNB, ComplementNB, BernoulliNB) являются:

## Тестовые вопросы по LinearDiscriminantAnalysis

1. LinearDiscriminantAnalysis (LDA) - это метод:
2. Какова основная цель LDA в контексте классификации?
3. Какое предположение делает LDA о распределении данных?
4. Что такое дискриминантные функции в LDA?
5. Какой параметр в LinearDiscriminantAnalysis отвечает за количество дискриминантных функций (компонент)?
6. Какой метод в LinearDiscriminantAnalysis используется для обучения модели?
7. Какой метод в LinearDiscriminantAnalysis используется для предсказания класса для новых данных?
8. Какой атрибут обученного объекта LinearDiscriminantAnalysis содержит коэффициенты дискриминантных функций?
9. LDA может быть использован для снижения размерности данных?
10. В каких случаях LDA может быть предпочтительнее PCA для снижения размерности?

# Тестовые вопросы по "Метод опорных векторов"

---

1. Метод опорных векторов (SVM) - это алгоритм:
2. Какова основная идея метода опорных векторов?
3. Что такое опорные векторы?
4. Что такое зазор (margin) в методе опорных векторов?
5. Что такое ядро (kernel) в методе опорных векторов?
6. Какие типы ядер обычно используются в методе опорных векторов?
7. Что такое параметр C в методе опорных векторов?
8. Какой метод в `sklearn.svm.SVC` используется для обучения модели?
9. Какой метод в `sklearn.svm.SVC` используется для предсказания класса для новых данных?
10. Метод опорных векторов чувствителен к масштабированию признаков?

## Тестовые вопросы по sklearn SVC

---

1. Какой класс в Scikit-Learn используется для реализации метода опорных векторов для классификации?
2. Какой параметр в `SVC` используется для выбора типа ядра?
3. Какие значения может принимать параметр `kernel` в `SVC`?
4. Какой параметр в `SVC` используется для управления степенью регуляризации?
5. Как влияет увеличение значения параметра `C` на сложность модели `SVC`?
6. Какой параметр в `SVC` используется для задания степени полинома при использовании полиномиального ядра?
7. Какой параметр в `SVC` используется для задания ширины гауссова ядра (RBF)?
8. Какой метод в `SVC` используется для обучения модели?
9. Какой метод в `SVC` используется для предсказания класса для новых данных?
10. Какой атрибут обученного объекта `SVC` содержит опорные векторы?

## Тестовые вопросы по SVC, NuSVC и LinearSVC

---

1. Какой из классов `SVC`, `NuSVC` и `LinearSVC` предназначен только для линейных ядер?
2. Какой параметр используется в `SVC` и `NuSVC` для управления степенью регуляризации, но отсутствует в `LinearSVC`?
3. Какой параметр в `NuSVC` контролирует долю ошибок обучения и опорных векторов?
4. Какой из классов, `SVC`, `NuSVC` или `LinearSVC`, обычно обучается быстрее на больших наборах данных с линейно разделимыми классами?

5. Какой из классов поддерживает больше типов ядер, включая нелинейные, такие как 'rbf' и 'poly'?
6. Какой параметр в `LinearSVC` используется для выбора типа регуляризации (L1 или L2)?
7. Какой из классов использует `liblinear` в качестве базовой библиотеки для оптимизации?
8. Какой из классов использует `libsvm` в качестве базовой библиотеки для оптимизации?
9. Какой из классов не поддерживает вероятностные оценки классов (`predict_proba`)?
10. Какой параметр в `SVC` и `NuSVC` позволяет использовать пользовательское ядро?