

ЦМВГН - квиз №2

* Indicates required question

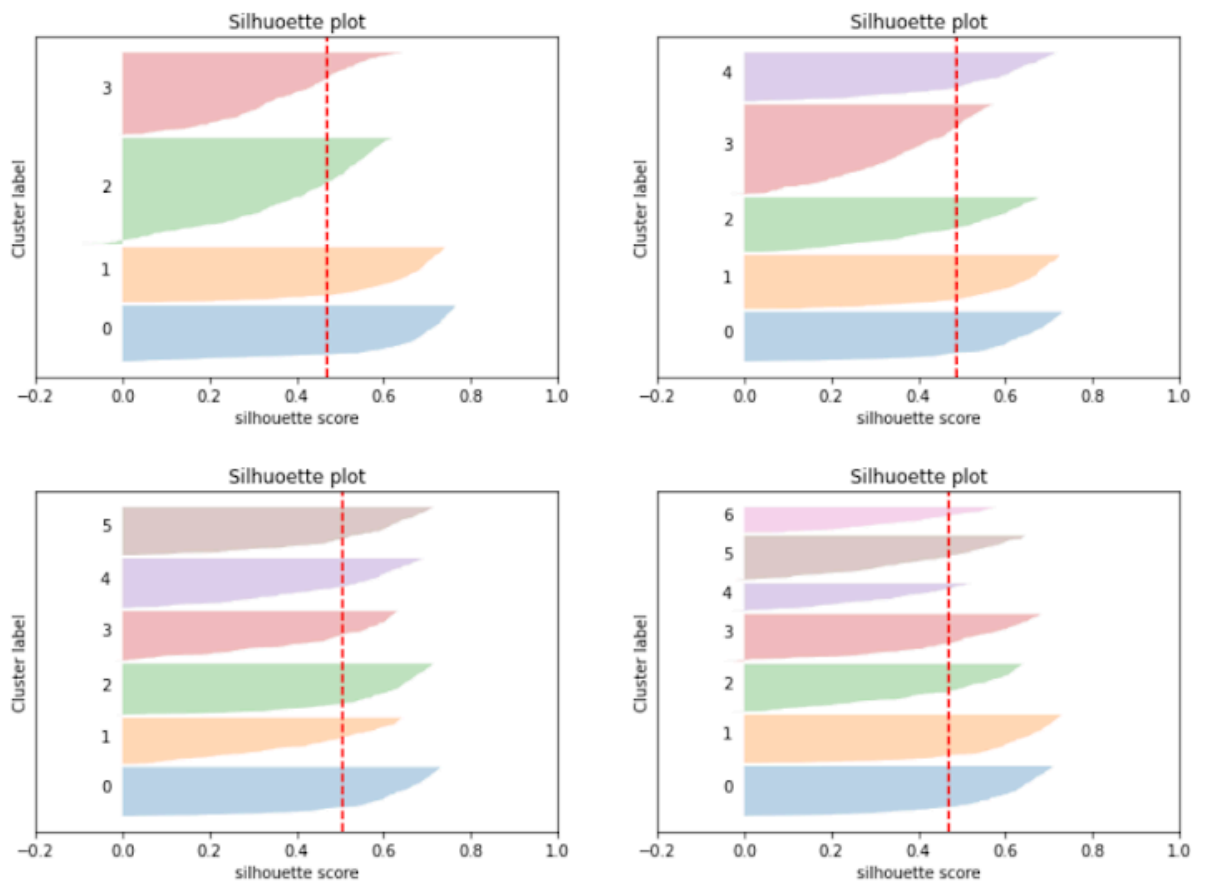
1. Email *

2. Ваше имя и фамилия *

Квиз

3. Определите оптимальное количество кластеров в данном случае, ориентируясь на силуэты:

* 1 point



Mark only one oval.

- ☐ 4
- ☐ 5
- ☐ 6
- ☐ 7

4. К вам попала чья-то оцифрованная домашняя библиотека. Никакой информации о книгах, кроме собственно текстов, не сохранилось. Вы хотите кластеризовать полученные тексты. Какие метрики качества кластеризации вы используете?

* 1 point

Check all that apply.

- ☐ Коэффициент силуэта
- ☐ Гомогенность и полнота
- ☐ Adjusted Rand Index
- ☐ V-мера

5. Посмотрите на предложенный фрагмент кода. Какую ошибку вы заметили? * 1 point

```
1 from sklearn.datasets import make_circles
2 from sklearn.cluster import DBSCAN
3 from sklearn.metrics import adjusted_rand_score
4
5
6 X, y = make_circles()
7 dbscan = DBSCAN()
8 dbscan.fit(X)
9 print(adjusted_rand_score(y, dbscan.predict(X)))
```

Mark only one oval.

- ☐ Здесь нельзя применять ARI
- ☐ У DBSCAN нет метода .predict()
- ☐ Не задано количество кластеров
- ☐ Предсказания происходят на всей выборке

6. Выберите **все** параметры, которые можно подобрать **при помощи GridSearch**. При необходимости можно и нужно обращаться к документации. * 1 point

Check all that apply.

- ☐ Коэффициент регуляризации в регрессиях Lasso и Ridge
- ☐ Коэффициенты/параметры в уравнении регрессии
- ☐ Стратегия оптимизации (solver)

7. Выберите **алгоритм ансамблирования моделей**, который позволит моделям учиться параллельно: * 1 point

Mark only one oval.

- ☐ Бэггинг
- ☐ Бутстрэп
- ☐ Бустинг

8. Сопоставьте метод ансамблирования и его интерпретацию с точки зрения смещения/дисперсии:

* 3 points

Mark only one oval per row.

	Бэггинг	Бустинг	Стекинг
Снижает смещение	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Снижает дисперсию	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Просто уменьшает ошибку	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

9. Выберите все утверждения, которые описывают градиентный и/или адаптивный бустинг. * 4 points

Внимание: утверждение может быть справедливым для обоих алгоритмов.

Check all that apply.

	Градиентный бустинг	Адаптивный бустинг
Начинает со "слабой" модели и пытается итеративно ее улучшить	<input type="checkbox"/>	<input type="checkbox"/>
Корректирует веса, присваивая больший вес объектам, на которых модель сработала неправильно	<input type="checkbox"/>	<input type="checkbox"/>
Делает новые предсказания, опираясь на разницу между векторами предсказаний;	<input type="checkbox"/>	<input type="checkbox"/>
Является линейной комбинацией входящих в него моделей	<input type="checkbox"/>	<input type="checkbox"/>

10. Выберите все верные утверждения о рандомном лесе: *

1 point

Check all that apply.

- ☐ Может состоять из любых моделей
- ☐ Выбирает подмножество не только из данных, но и из признаков
- ☐ Может решать задачи классификации и регрессии
- ☐ Использует неглубокие деревья

11. Выберите все **методы снижения размерности**, при помощи которых * 1 point
можно извлечь значимые кластеры из данных.

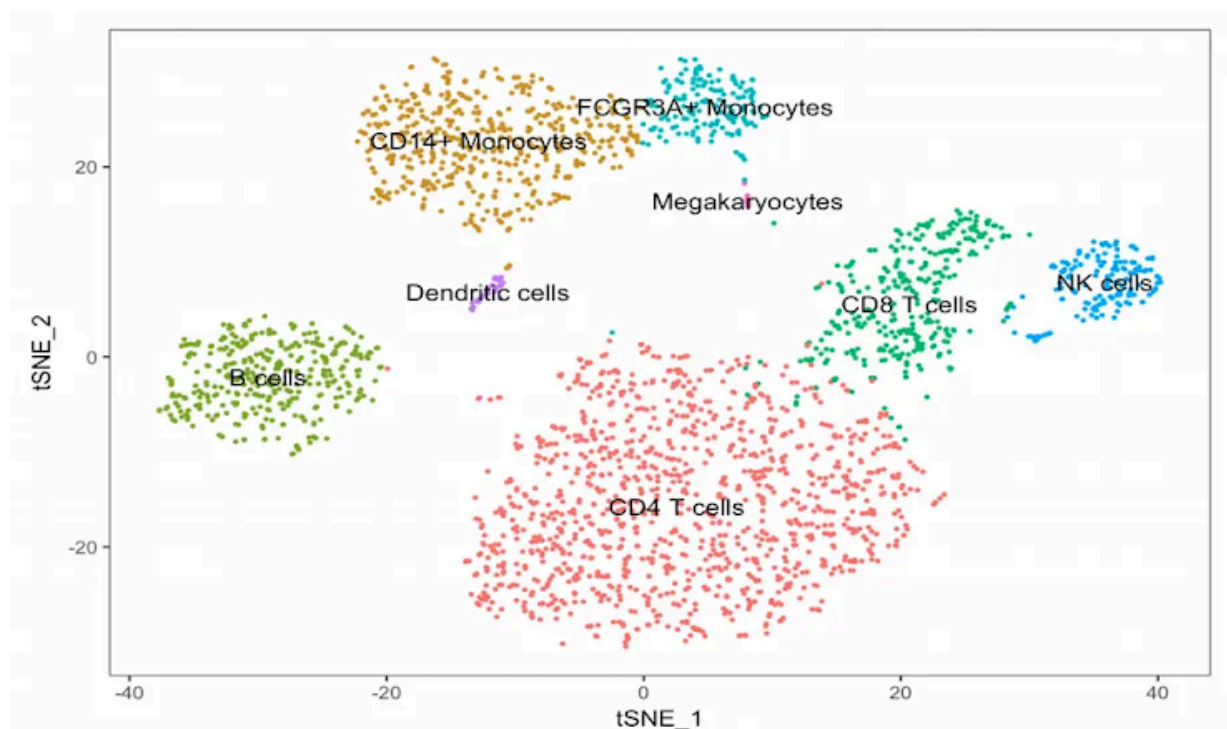
Check all that apply.

- ☐ Latent Dirichlet Allocation
- ☐ K Means
- ☐ t-SNE

12. Данная визуализация была создана при помощи tSNE (автор - Николай Осколков: <https://towardsdatascience.com/how-to-tune-hyperparameters-of-tsne-7c0596a18868>).

* 1 point

Можем ли мы с уверенностью утверждать, что кластеры NK cells и B cells находятся далеко друг от друга в реальном (не сжатом) пространстве признаков?



Mark only one oval.

- ☐ Да
- ☐ Нет

This content is neither created nor endorsed by Google.

Google Forms

