



МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«МОСКОВСКИЙ АВИАЦИОННЫЙ ИНСТИТУТ
(национальный исследовательский университет)»
Институт №12 «Аэрокосмические наукоемкие технологии и производства»
Кафедра «Технология испытаний и эксплуатации»

Выпускная квалификационная работа магистра на тему:

Разработка алгоритмов машинного обучения для оценки дизайна веб-сайтов

Автор работы: студент гр. Т12О-201М-20, Меркулов Лев Владимирович

Руководитель: к.т.н., доцент каф. 806 МАИ, Ухов Пётр Александрович

Москва
2022

Актуальность темы

- **48%** людей считают дизайн веб-сайта важнейшим пунктом в принятии решения о доверии к бизнесу.
- Пользователи уходят с сайта, если за 10 секунд не получили основную информацию об услуге.
- Визуально привлекательный сайт способен увеличить регистрации и продажи.
- **38%** людей перестанут изучать сайт, если контент покажется им непривлекательным;
- Экспертная оценка качества дизайна субъективна и занимает много времени.
- На сегодняшний день не существует популярных инструментов автоматизации оценивания дизайна.



«На что вы обращаете внимание перед первой покупкой на сайте?»
Источник: Econsultancy

Постановка проблемы

- Полезность разработки: продукт или дополнительная функция продукта в сфере веб-дизайна для оценки качества веб-сайта с помощью алгоритмов машинного обучения. Предполагается, что это обеспечит поток новых пользователей, повысит вероятность купить продукт и рекомендовать его знакомым.
- Кто потребитель: веб-дизайнеры-новички, с годом опыта и менее, не обладающие большой насмотренностью, а также заказчики дизайн-проектов.
- С какой целью будет использовать: получить оценку и рекомендации по доработке дизайна. Публикация работы на специальных ресурсах с целью получить оценку от пользователей может затянуться и доставить дискомфорт - начинающие дизайнеры могут стесняться публиковать свои работы.

Цель и задачи работы

Цель работы заключается в автоматизации процесса оценки качества дизайна веб-сайта с помощью моделей машинного обучения.

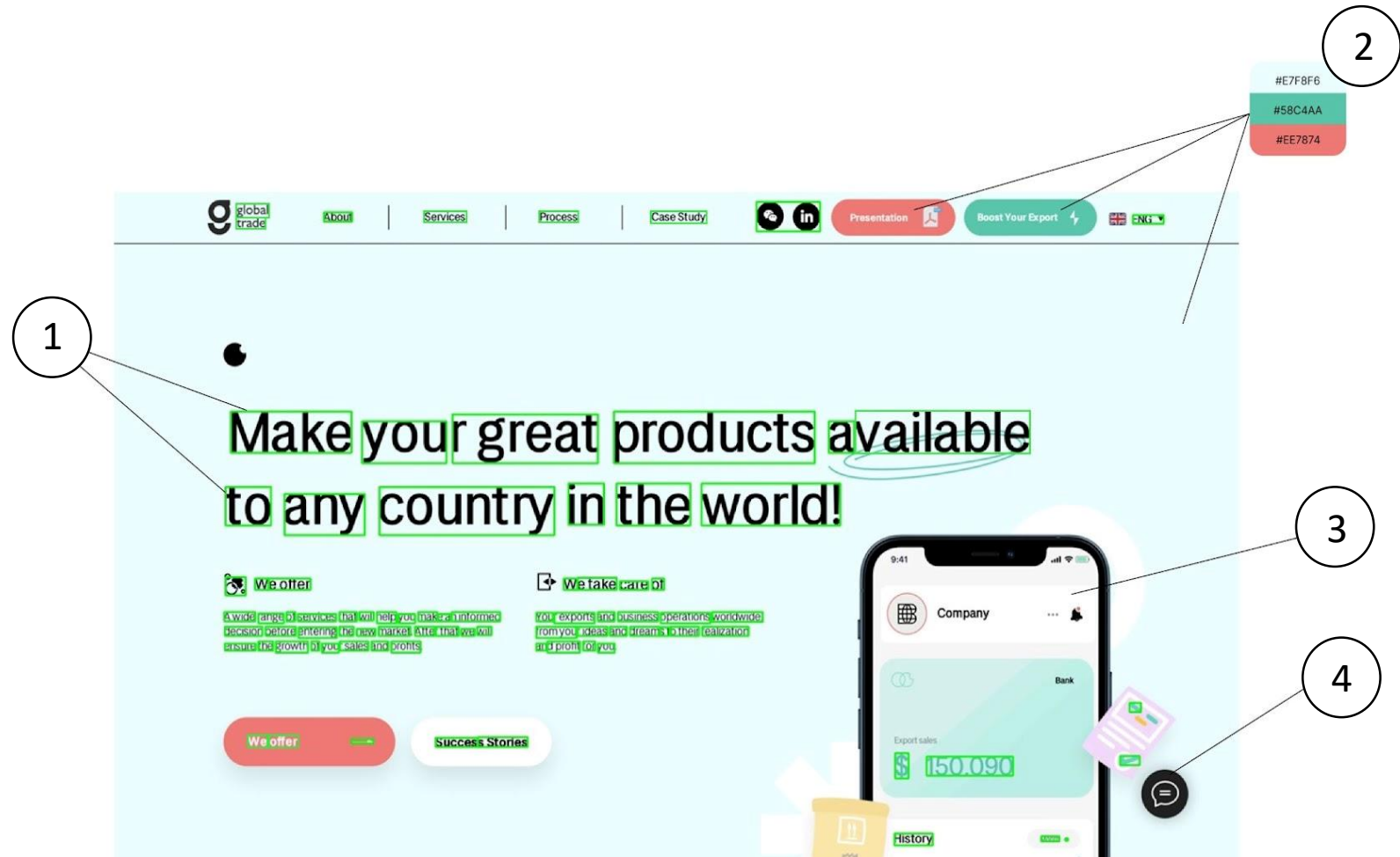
Для этого были поставлены следующие задачи:

- Определить признаки изображений, связанные с качеством дизайна;
- Собрать набор данных и провести их анализ;
- Выбрать алгоритмы машинного обучения для решения задачи классификации;
- Выбрать метрики для оценки эффективности моделей;
- Обучить модели и сравнить их результаты по метрикам;
- Отобрать наиболее важные признаки;
- Обучить модель на отобранных признаках.

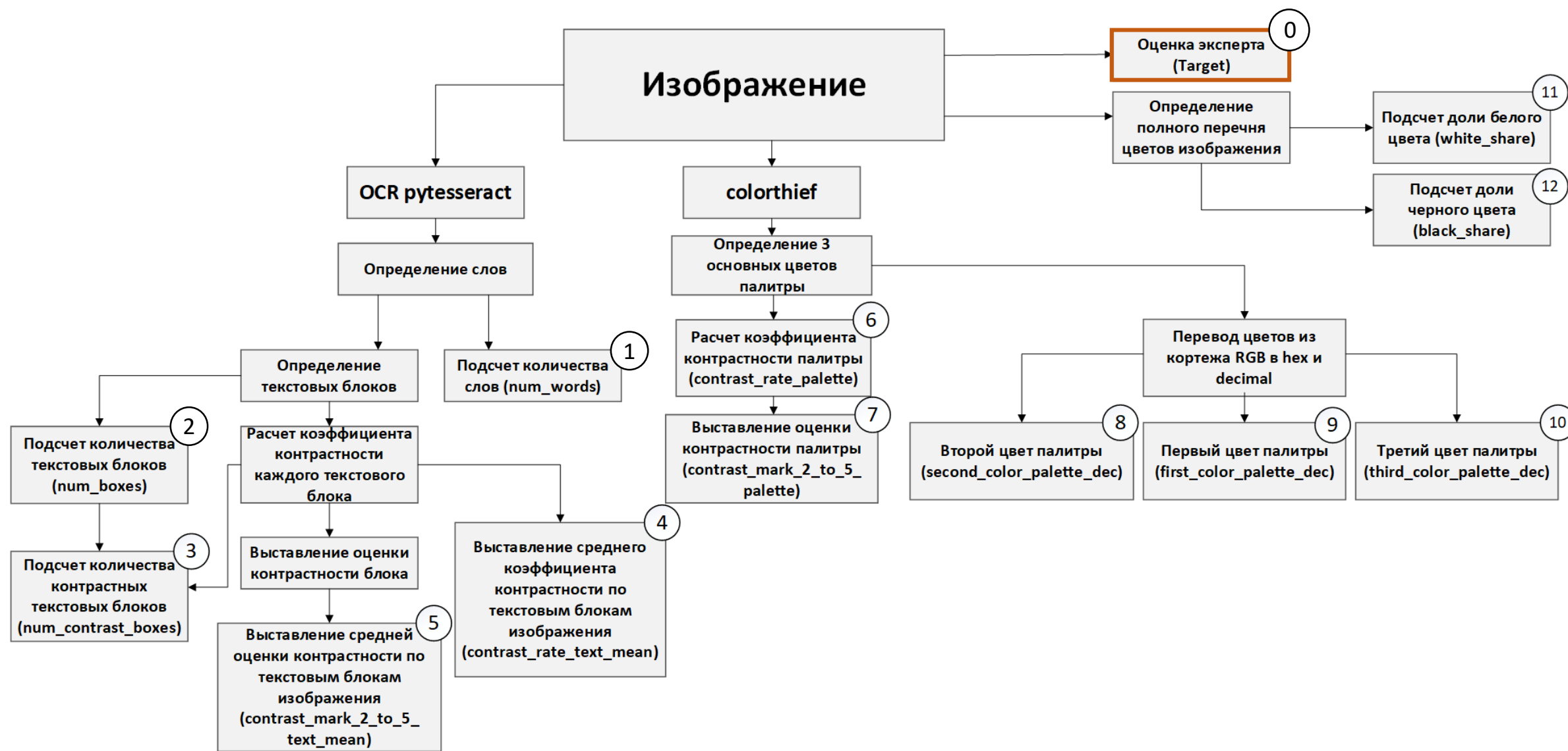
Критерии для оценки качества дизайна

Контрольная группа из 5 экспертов в веб-дизайне выделила следующие основные критерии:

- Слова на лендинге (общее число и количество контрастных);
- Цветовая палитра изображения (основные цвета, оценка сочетаемости);
- Доля белого цвета;
- Доля черного цвета.



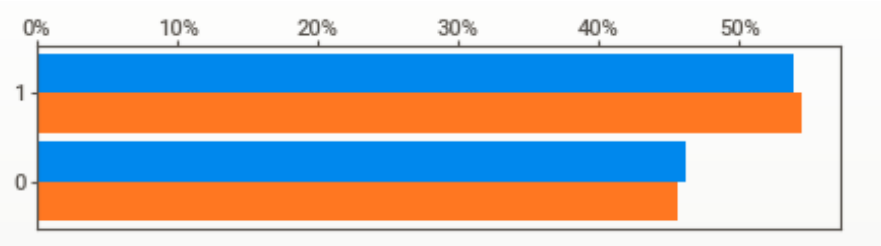
Выбранные признаки





Перечень выбранных признаков

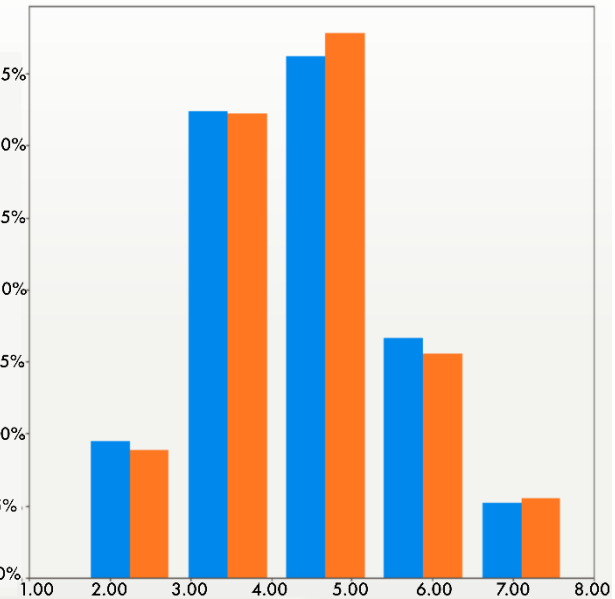
1. **Количество слов** (num_words) - число боксов, определенных как текст;
2. **Количество текстовых групп** (num_boxes) - число боксов, которые соответствуют заданному в программе уровню “уверенности”;
3. **Количество контрастных текстовых групп** (num_contrast_boxes) - количество текстовых групп, превышающих минимальный уровень контрастности;
4. **Коэффициент контрастности** (contrast_rate_text) - коэффициент контрастности текста;
5. **Оценка контрастности** (contrast_mark_2_to_5_text) - оценка соответствия контрастности требованиям доступности контента на веб-странице;
6. **Палитра из трех основных цветов дизайна**;
7. **Коэффициент контрастности палитры** (contrast_rate_palette) - среднее между контрастностью трех основных цветов палитры;
8. **Оценка контрастности палитры** (contrast_mark_2_to_5_palette) - оценка соответствия контрастности палитры требованиям доступности контента на веб-странице;
9. **Доля белого и черного цветов** (white_share, black_share) - поскольку черный и белый цвета не входят в цветовой круг, их процентные доли в изображении считаются отдельно.

Анализ данных

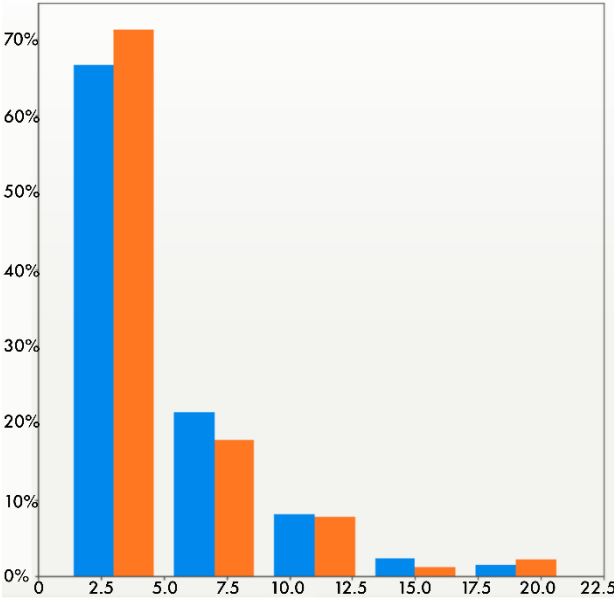


Распределение
целевой переменной

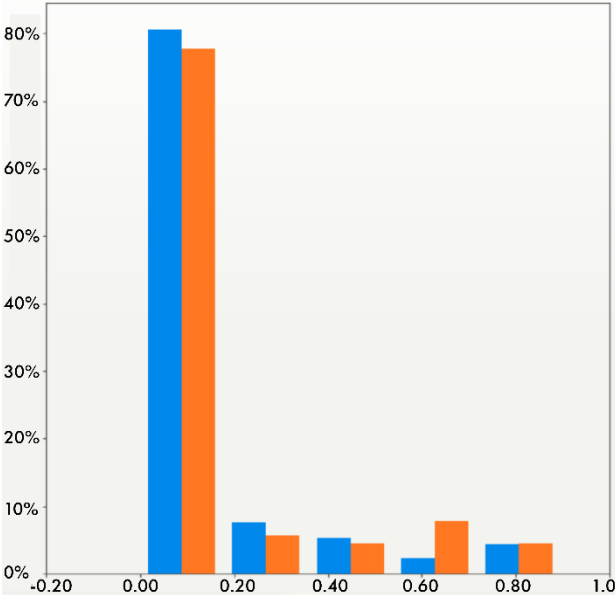
 - обучающая выборка
 - тестовая выборка



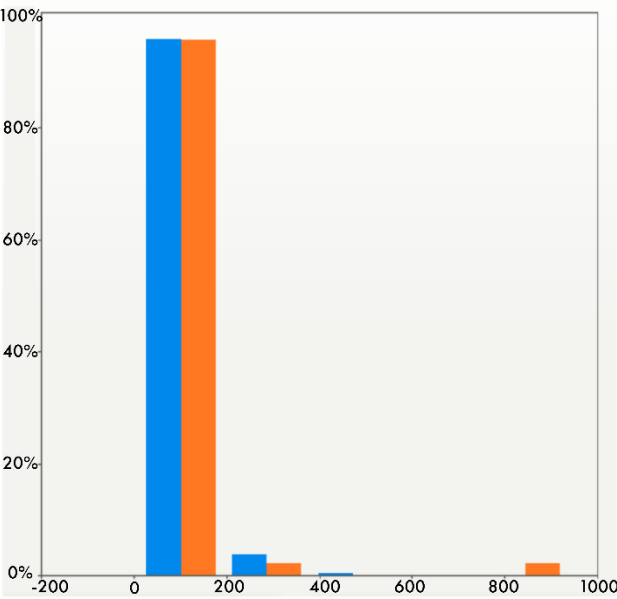
Контрастность палитры



Контрастность текста



Доля белого



Количество слов

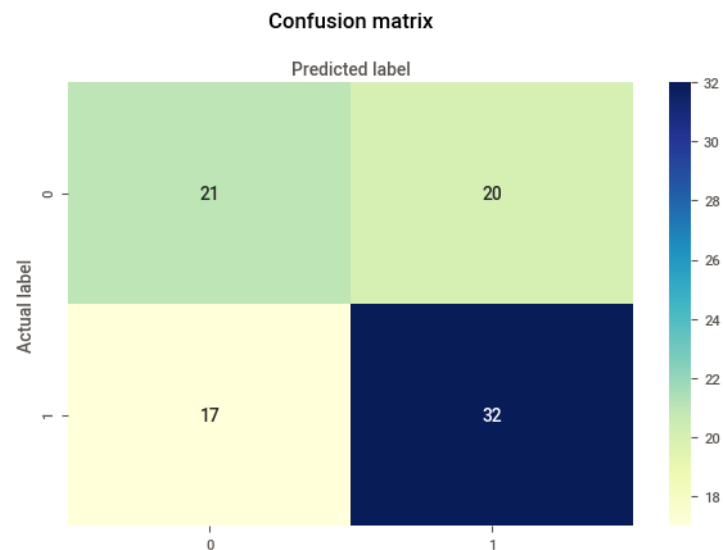
Эффективность моделей

Модель	Accuracy	Recall	Precision	f1-score	Average precision score
Логистическая регрессия - стандартизация	0,57	0,69	0,59	0,64	0,57
Логистическая регрессия - нормализация	0,52	0,78	0,54	0,64	0,54
Логистическая регрессия - бинаризация	0,59	0,84	0,59	0,69	0,58
Дерево решений	0,56	0,57	0,60	0,58	0,57
Случайный лес	0,57	0,63	0,60	0,61	0,58
XGBoost (порог = 0.5)	0,59	0,65	0,62	0,63	0,59
XGBoost (порог = 0.2)	0,67	0,82	0,66	0,73	0,64

Оцениваем по Precision, поскольку она считает долю верных предсказаний класса 1

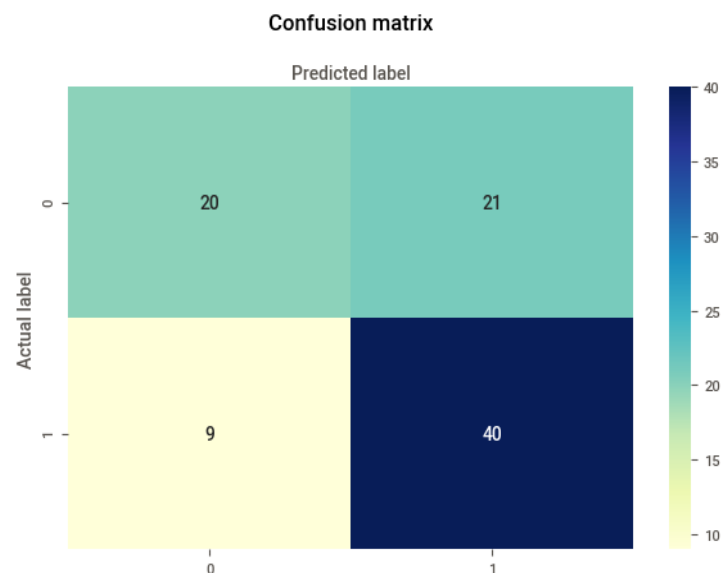
Также по Average precision, поскольку она характеризует среднее значение точности, рассчитанное для каждого порога

Эффективность моделей



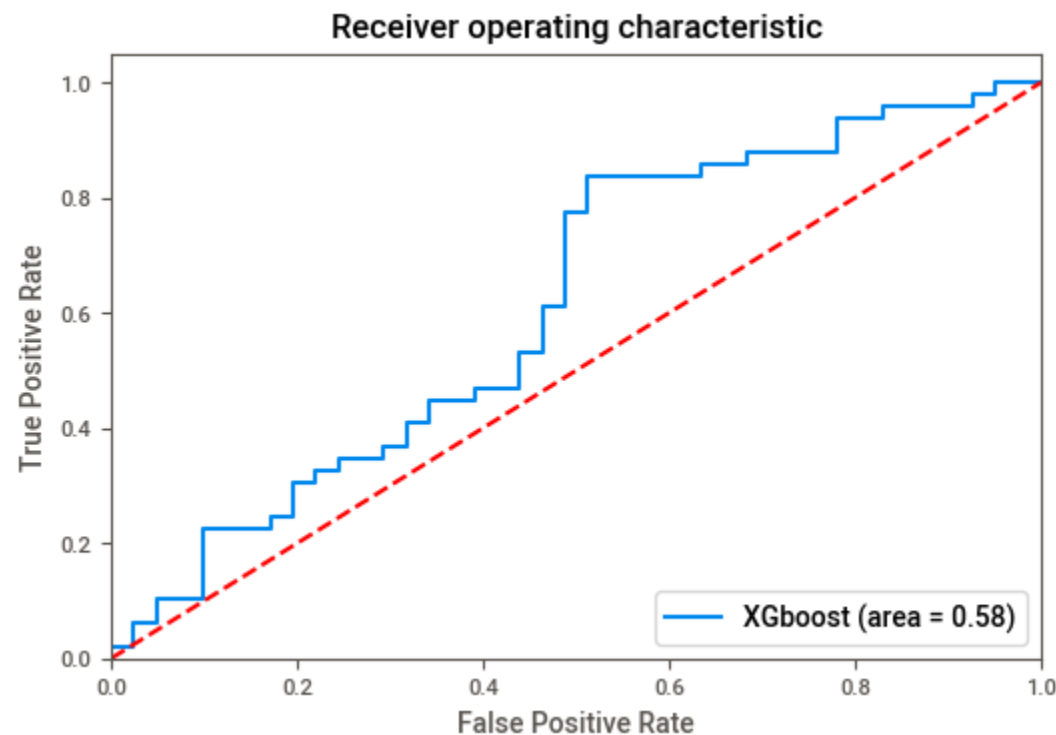
Матрица ошибок
XGBoost (порог = 0.5)

Синий квадрат – правильно
определенный
качественный дизайн



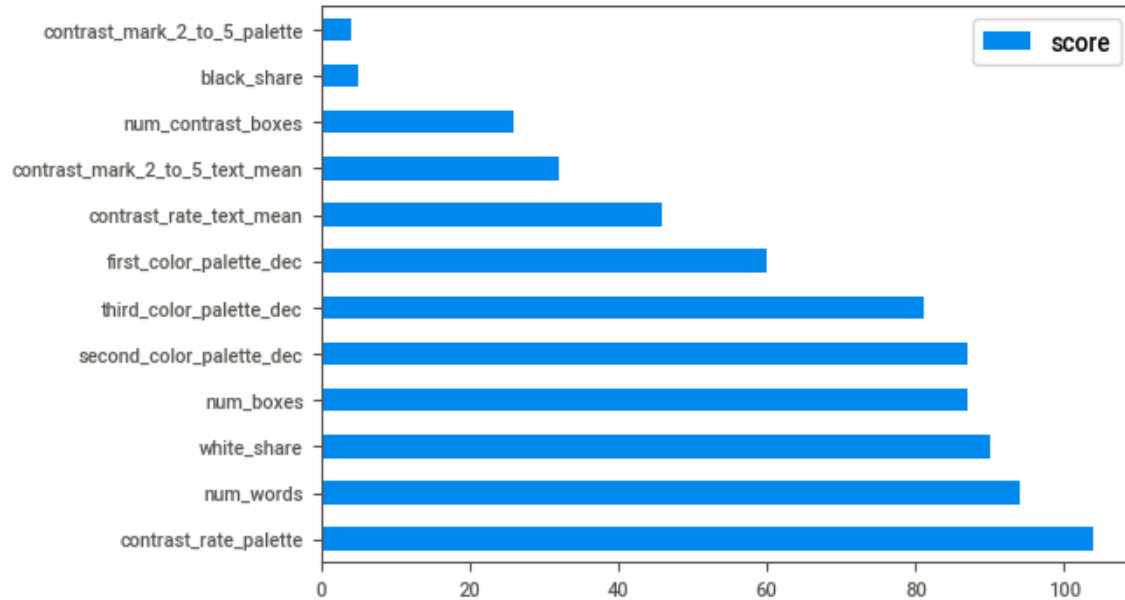
Матрица ошибок
XGBoost (порог = 0.2)

Синий квадрат – правильно
определенный
качественный дизайн

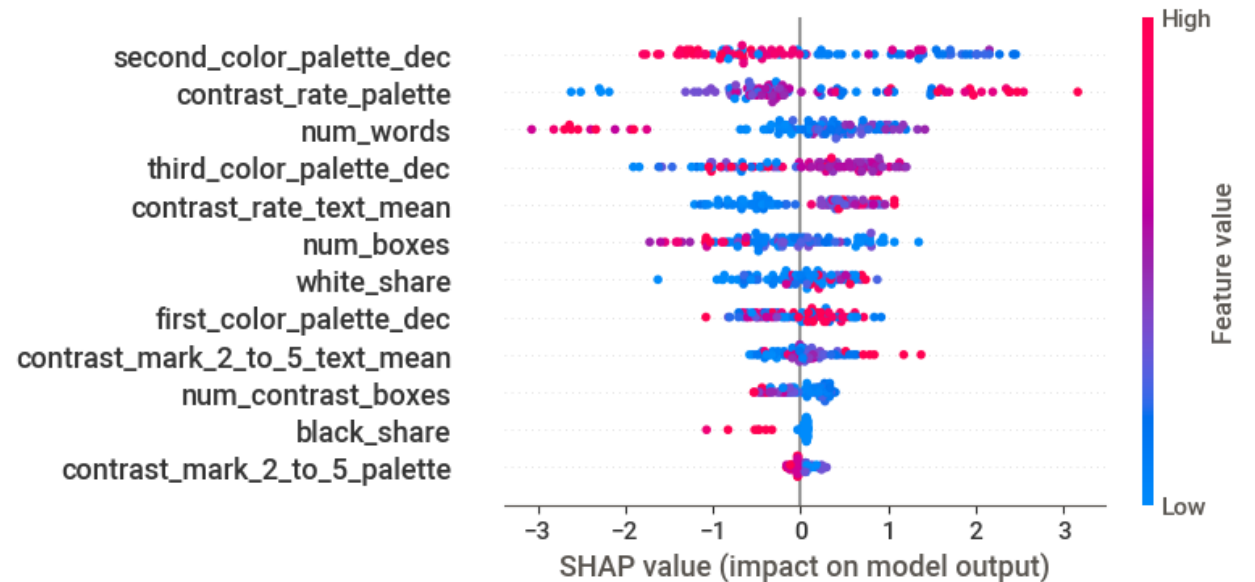


ROC-кривая XGBoost (порог = 0.5)
Чем лучше алгоритм, тем ближе площадь
под кривой (AUC ROC) к 1

Отбор признаков по важности



Перечень признаков в порядке возрастания их влияния на предсказание модели



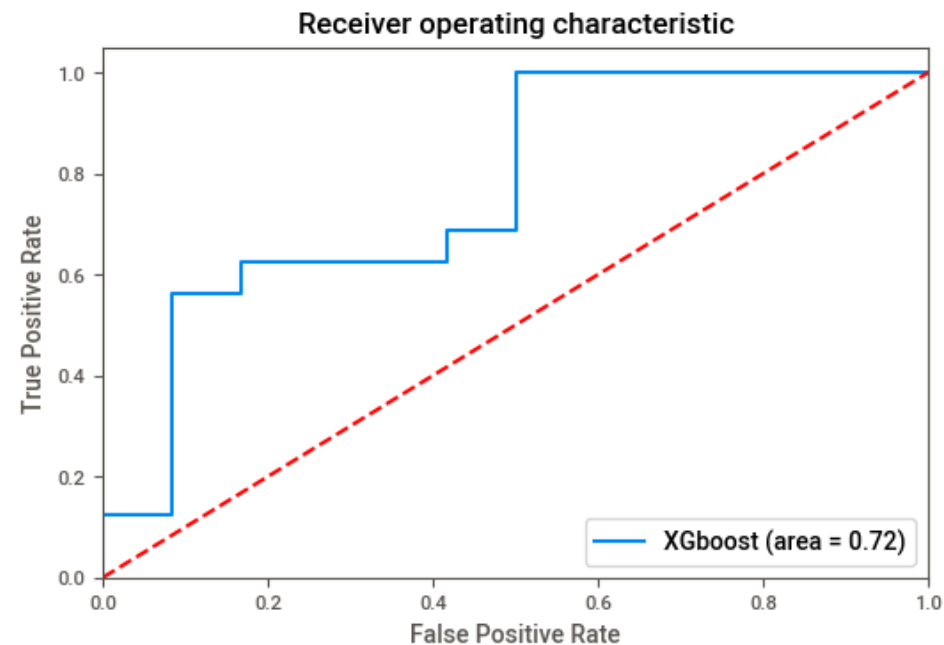
Распределение влияния каждого признака на предсказание модели.

- Красный и синий цвета – высокое и низкое значение признака.
- Значения слева от центральной вертикальной линии — это некачественный дизайн, справа — качественный

Результаты обучения по отобранным признакам

	XG Boost (все признаки)	XG Boost (порог = 0.5)	XG Boost (порог = 0.4)
Precision	0,59	0,71	0,70
Recall	0,65	0,94	1,00
F1-score	0,62	0,81	0,82
Accuracy	0,63	0,75	0,75
Average Precision Score	0,59	0,71	0,70

Результаты модели после отбора признаков.
В первом столбце представлены показатели предыдущей модели.



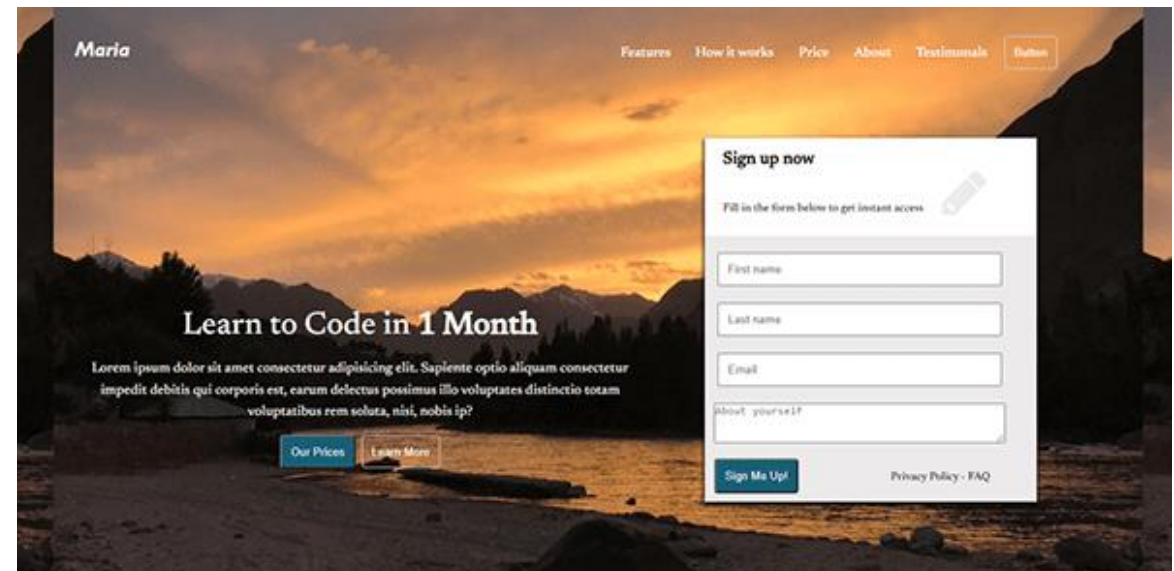
ROC-кривая.
Чем лучше алгоритм, тем ближе площадь под кривой (AUC ROC) к 1.

Результаты. Некачественный дизайн

	num_words	num_boxes	num_contrast_boxes	contrast_rate_text_mean	contrast_mark_2 to_5_text_mean	contrast_rate palette	color palette
Пример 1	100	22	2	2,347	2,272	3,693	(12, 15, 14) (192, 156, 154) (21, 172, 87)
Пример 2	22	5	1	5,026	2,600	3,096	(195, 151, 113) (31, 26, 25) (78, 55, 42)



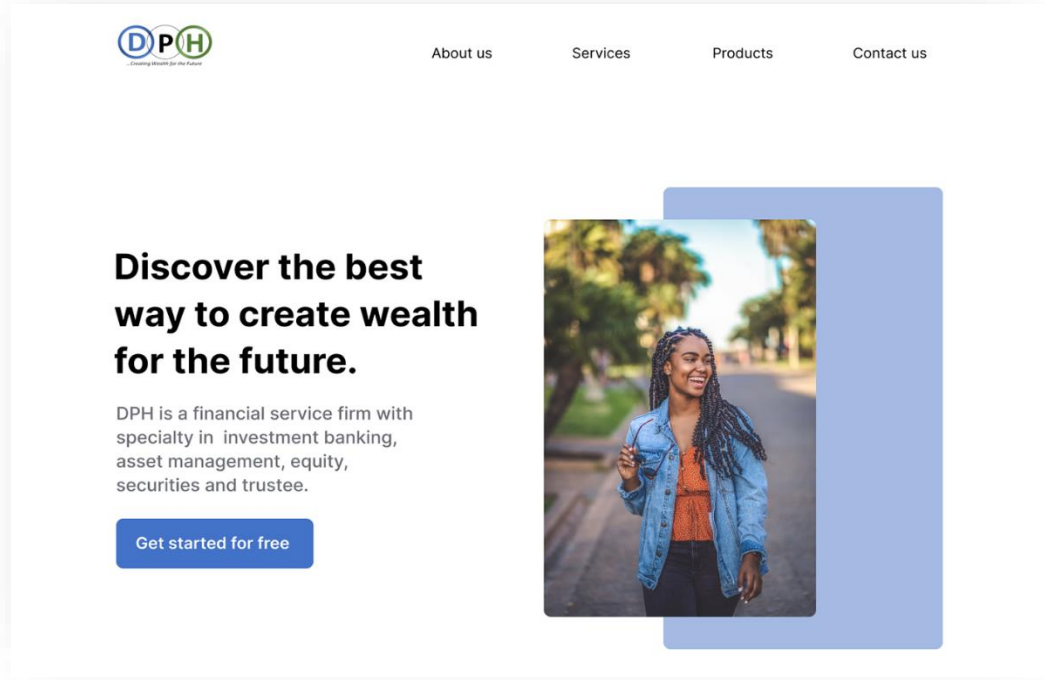
Пример 1



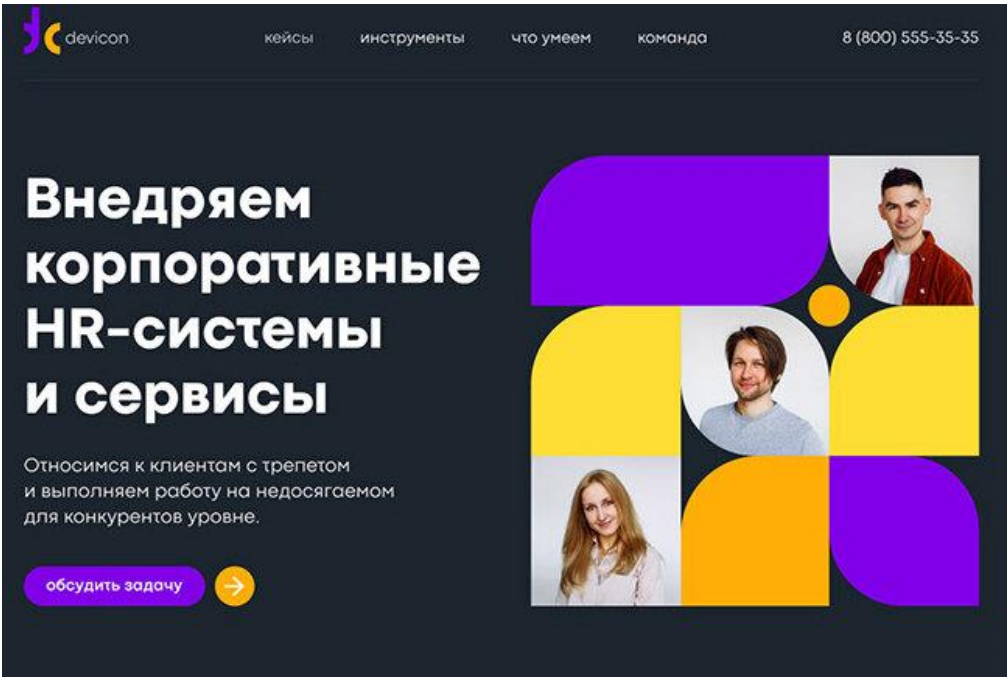
Пример 2

Результаты. Качественный дизайн

	num_words	num_boxes	num_contrast_boxes	contrast_rate_text_mean	contrast_mark_2 to_5_text_mean	contrast_rate palette	color palette
Пример 1	60	38	28	9,454	3,710	6,037	(156, 166, 179) (46, 61, 86) (94, 76, 70)
Пример 2	43	5	3	9,754	3,800	6,273	(43, 32, 69) (230, 204, 126) (139, 148, 154)



Пример 1



Пример 2

Итог работы

- Составлен перечень признаков, связанных с качеством дизайна;
- Собран сбалансированный набор данных из 300 изображений;
- Выбраны алгоритмы машинного обучения для решения задачи классификации;
- Выбраны метрики для оценки эффективности модели;
- Выполнено сравнение результатов моделей;
- Отобраны наиболее важные признаки;
- В дальнейшем планируется увеличить набор данных и добавить новые признаки.

Подтверждена гипотеза, что искусственный интеллект подходит для решения задачи оценки качества веб-дизайна.

Спасибо за внимание!

**Разработка алгоритмов машинного обучения
для оценки дизайна веб-сайтов**

Меркулов Лев Владимирович, Т12О-201М-20