

# Требования к защите проектов 1

## Разбалловка

Максимальная оценка за проект на защите: 15 баллов. Она складывается из оценки за доклад: максимум 10, и оценки за ответы на вопросы: максимум 5.

## Порядок сдачи

Группы выступают согласно полученному заранее расписанию (время примерное), допускается подойти за 10 мин до своего времени. Время на доклад: 8 + 2 мин (после 8 минут нужно переходить к выводам). Время на вопросы: 5 мин.

## Требования к докладу

В докладе должны быть чётко сформулированы:

1. Цель исследования (её вы определяете сами, это цель **всего** проекта)
2. Задачи (шаги к достижению цели и какие из них были вынесены к первой защите)
3. Выводы (какие из задач получилось закрыть, с какими результатами и насколько это приблизило вас к цели)

Докладывать может один или несколько докладчиков, но в докладе должны быть обозначены зоны ответственности всех участников группы. Отвечать на вопросы может не только докладчик, но и другие участники его группы. Один из участников группы должен фиксировать вопросы и после защиты прислать список "белых мест", подсвеченных группе.

## Требования к содержанию

В вашем исследовании должно быть представлено:

1. Описание датасета (с какими данными вы работаете, какие есть фичи).
2. Вы должны выбрать два параметра: непрерывный и дискретный, и сформулировать для них постановку задачи регрессии и классификации соответственно (рекомендуемые параметры указаны в таблице вашего проекта, но вы можете выбрать другие).
3. Произвести предобработку данных (для классификации и регрессии она может быть разная). Желательно хотя бы применить базовые методы: стандартизировать, заполнить отсутствующие данные, почистить от выбросов. Можете попробовать добавлять новые фичи, преобразовывать категориальные фичи в численные, трансформировать данные и всё что угодно, что вы придумаете (главное обосновывайте, зачем вы это делаете).
4. Решение поставленных вами задач регрессии и классификации с использованием освоенных вами ранее инструментов (Логистическая регрессия, LDA, QDA, линейная модель, можно использовать и другие регрессии из курса АнДана).
5. В разделе с исследованием данных, предобработкой или диагностикой модели классификации (в зависимости от вашей задачи) должна быть кластеризация и попытка понижения размерности.

## Требования к оформлению

1. Белый фон и чёрный текст

2. Заголовки на всех слайдах

3. Нумерацию слайдов

4. Графики должны иметь названия, подписи осей, единицы измерения, сетку, легенду

Чтобы улучшить восприятие вашей презентации, постарайтесь выдержать её в едином стиле.

**Важно!**

Помните, что когда вы формулируете задачу, формулируйте также критерии её выполнения, например, если вы хотите построить лучший классификатор: опишите, какой классификатор вы будете считать лучшим и почему.

НЕ нужно делать просто перебор всех методов в исследовании, вы можете делать это в своих черновиках, но в докладе нужно представить исследование, т.е. шаги работы с логическими переходами между ними.

Ну и не забывайте про АнДан, вы можете пользоваться всем что уже знаете оттуда, но не используйте методы, которые ещё не разбирались в данном курсе, с ними вы будете работать на второй защите.