

МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ

УЧРЕЖДЕНИЕ ОБРАЗОВАНИЯ  
«БРЕСТСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

ФАКУЛЬТЕТ ЭЛЕКТРОННО-ИНФОРМАЦИОННЫХ СИСТЕМ

Кафедра интеллектуальных информационных технологий

## Отчет по лабораторной работе №5

Специальность ИИ-23

Выполнил

Медведь П.В.,

студент группы ИИ-23

Проверила

Андренко К.В.,  
преподаватель стажёр  
кафедры ИИТ,

«\_\_» \_\_\_\_ 2025 г.

Брест 2025

**Цель:** На практике сравнить работу нескольких алгоритмов одиночного дерева решений, случайного леса и бустинга для деревьев решений.

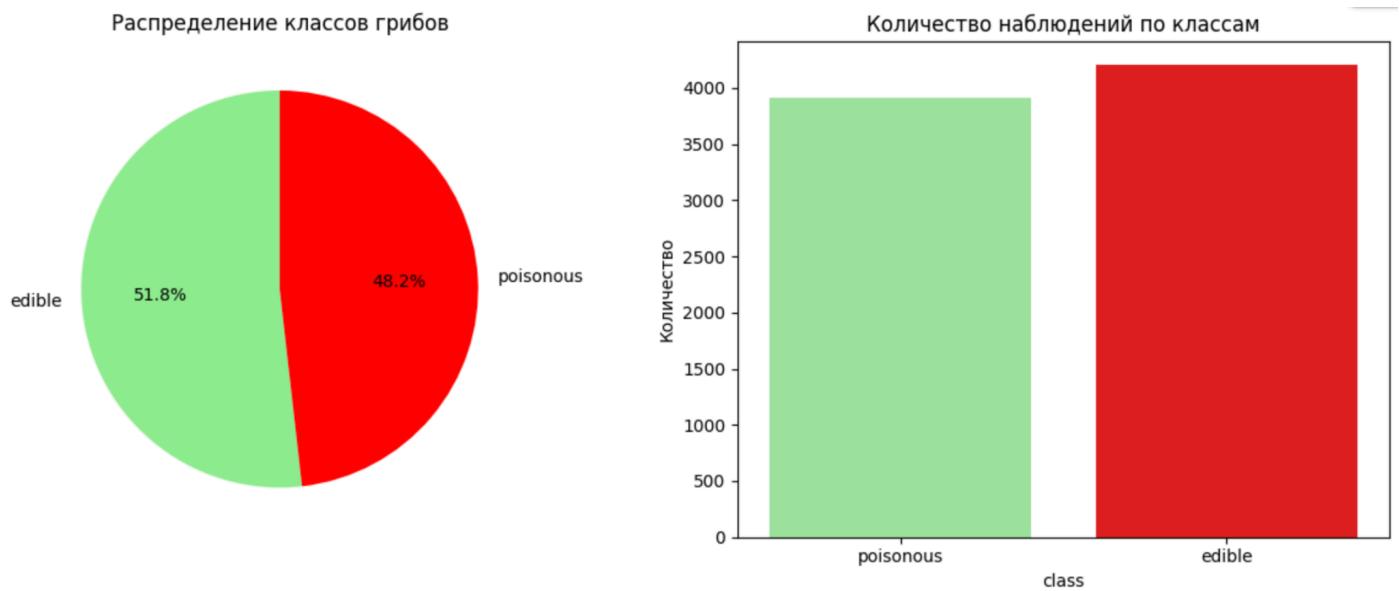
### Задачи:

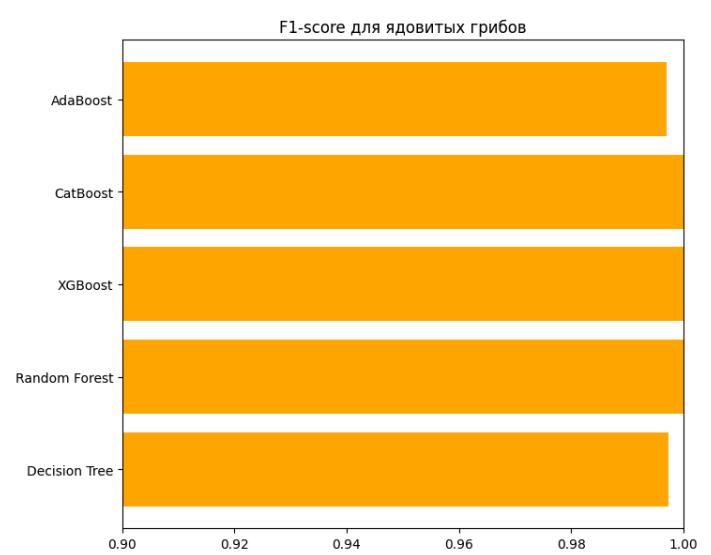
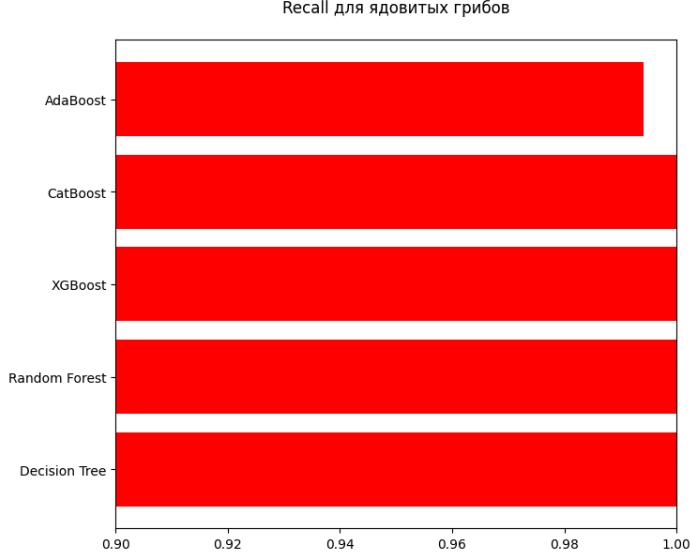
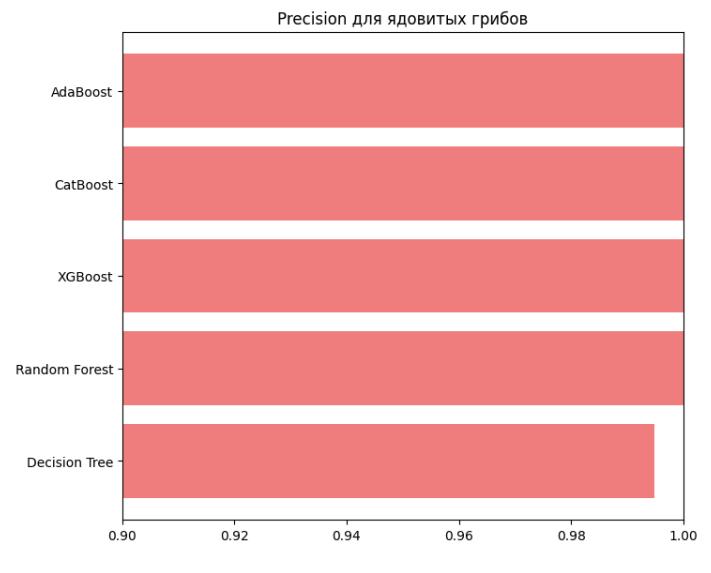
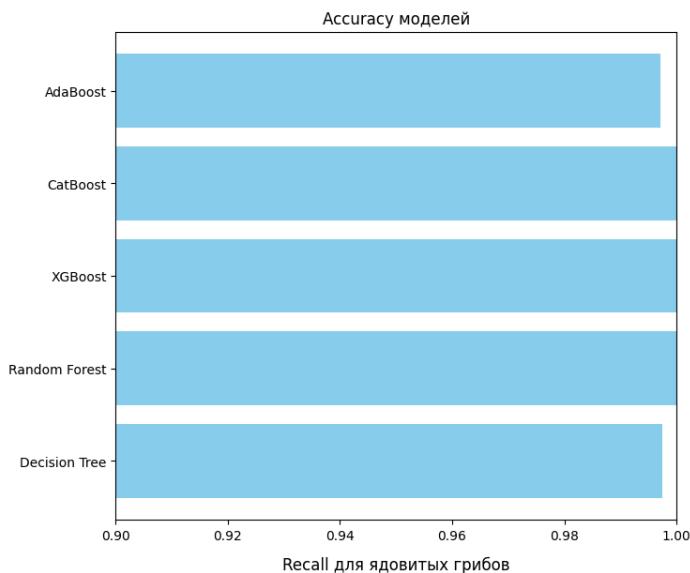
1. Загрузить датасет по варианту;
2. Разделить данные на обучающую и тестовую выборки;
3. Обучить на обучающей выборке одиночное дерево, случайный лес и реализовать бустинг для решающих деревьев (AdaBoost, CatBoost, XGBoost);
4. Оценить точность каждой модели на тестовой выборке;
5. Сравнить результаты, сделать выводы о применимости каждого метода для данного набора данных.

### Вариант 5

- Mushroom Classification
- Определить, является ли гриб ядовитым или съедобным
- **Задания:**
  1. Загрузите данные и преобразуйте все категориальные признаки в числовые (например, с помощью One-Hot Encoding);
  2. Разделите данные на обучающую и тестовую части;
  3. Обучить на обучающей выборке одиночное дерево, случайный лес и реализовать бустинг для решающих деревьев (AdaBoost, CatBoost, XGBoost);
  4. Рассчитайте точность и полноту (precision и recall) для класса "ядовитый";
  5. Сделайте вывод о том, какой классификатор лучше всего справляется с этой задачей, где цена ошибки очень высока.

### Результат:



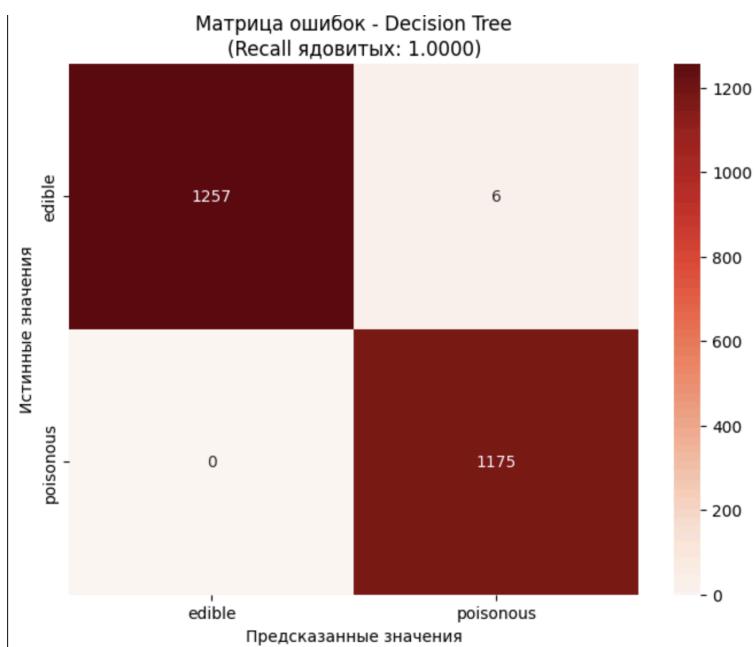


## ЛУЧШАЯ МОДЕЛЬ: Decision Tree

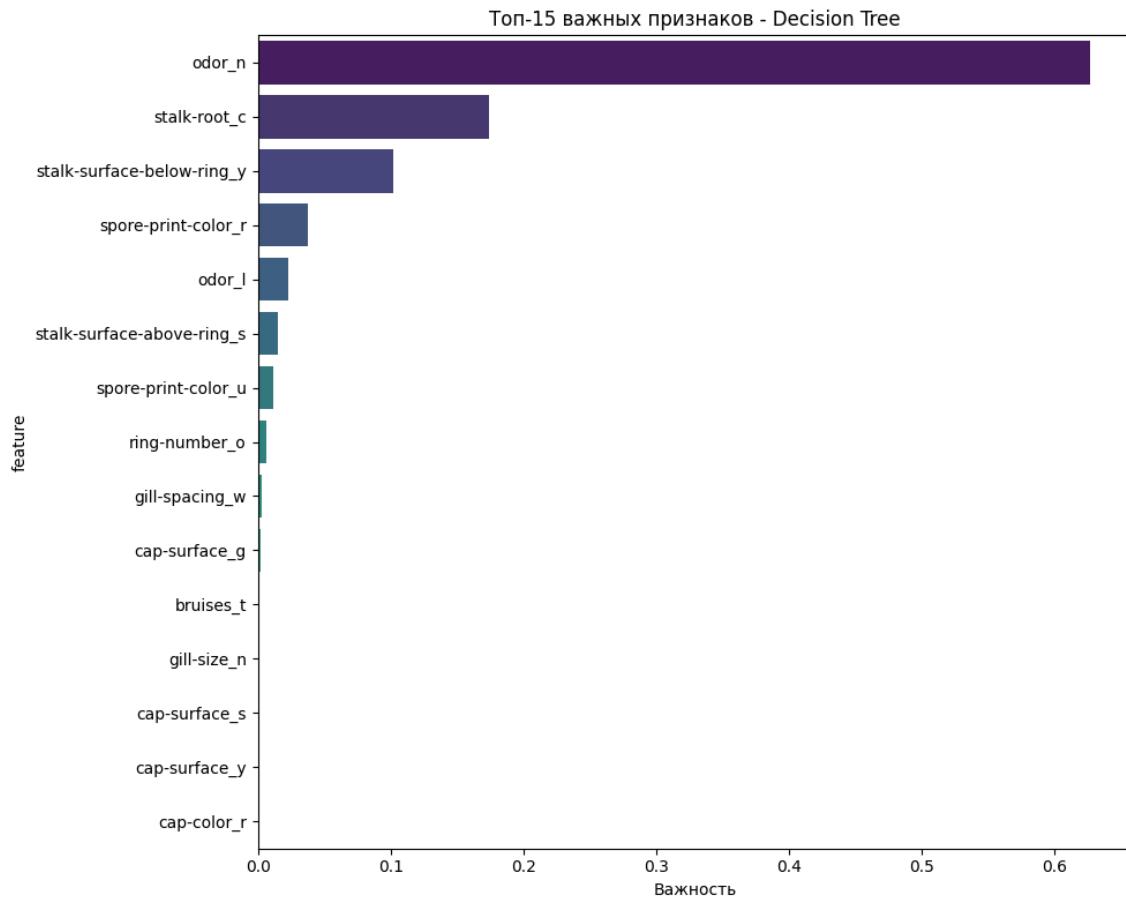
Recall для ядовитых: 1.0000

Детальный отчет классификации для Decision Tree:

	precision	recall	f1-score	support
edible (съедобные)	1.00	1.00	1.00	1263
poisonous (ядовитые)	0.99	1.00	1.00	1175
accuracy		1.00	2438	
macro avg	1.00	1.00	1.00	2438
weighted avg	1.00	1.00	1.00	2438



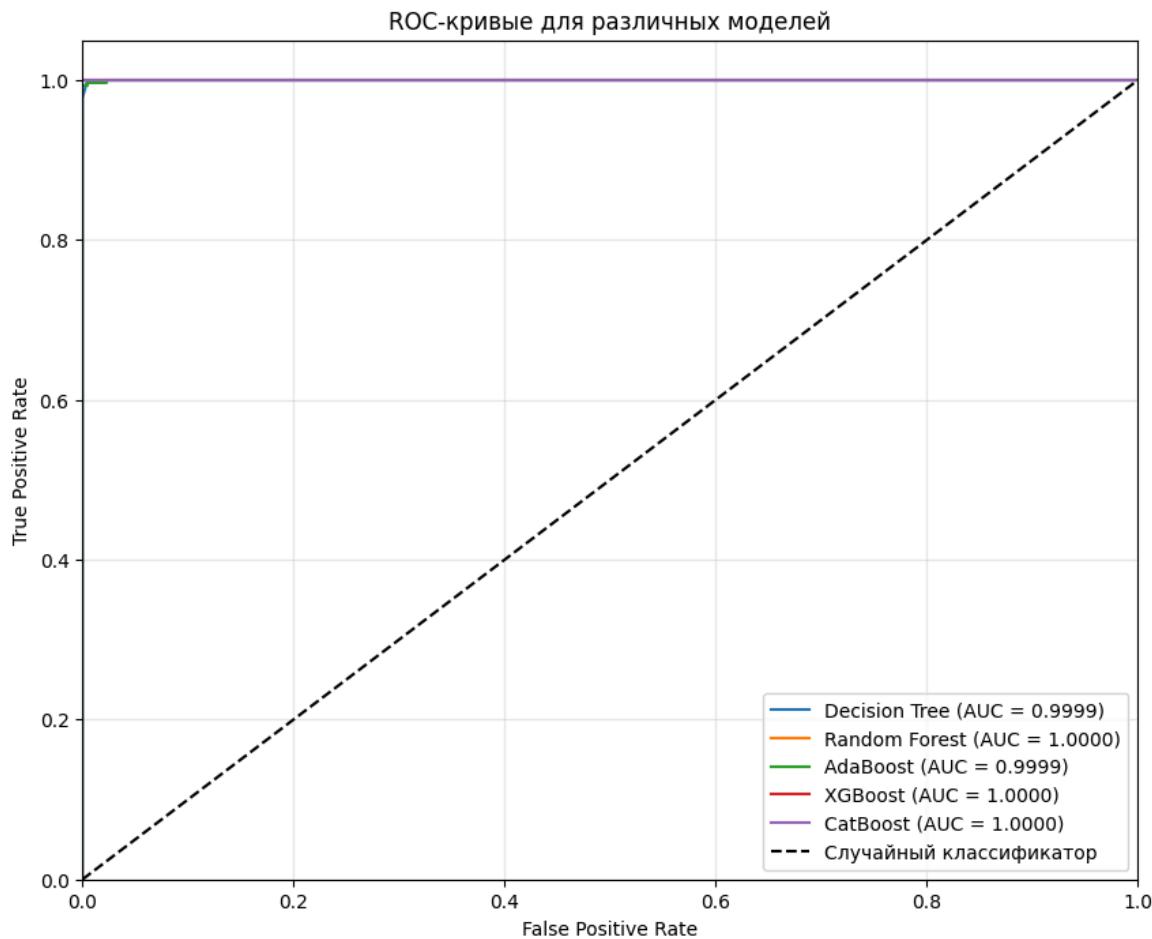
## Анализ важности признаков для Decision Tree



Топ-10 самых важных признаков:

feature	importance
odor_n	0.626606
stalk-root_c	0.174365
stalk-surface-below-ring_y	0.101892
spore-print-color_r	0.037238
odor_l	0.022606
stalk-surface-above-ring_s	0.014716
spore-print-color_u	0.011163
ring-number_o	0.006040
gill-spacing_w	0.002533
cap-surface_g	0.002130

ROC-кривые для моделей:



## Вывод:

Лучшая модель: Decision Tree

Recall для ядовитых: 1.0000

Анализ ошибок лучшей модели:

Правильно предсказано съедобных: 1257

Ложные срабатывания (съедобный → ядовитый): 6

Пропущенные ядовитые (ядовитый → съедобный): 0

Правильно предсказано ядовитых: 1175

Критические ошибки (пропущенные ядовитые): 0

Это означает, что 0 ядовитых гриба были ошибочно классифицированы как съедобные