Министерство цифрового развития

Федеральное государственное бюджетное образовательное учреждение высшего образования

«Сибирский государственный университет телекоммуникаций и информатики»

(СибГУТИ)

Кафедра прикладной математики и кибернетики

Отчёт

по лабораторной работе № 3 «Классификация методом дерева решений»

Выполнил:

студент группы ИП-312

Дорогин Н.С.

Работу проверил: старший преподаватель кафедры ПМиК Дементьева К.И.

Новосибирск 2025 г.

# **Введение (задание)**

Цель: освоить практическое применение метода решающего дерева для задач

классификации. Исследовать влияние гиперпараметров на качество

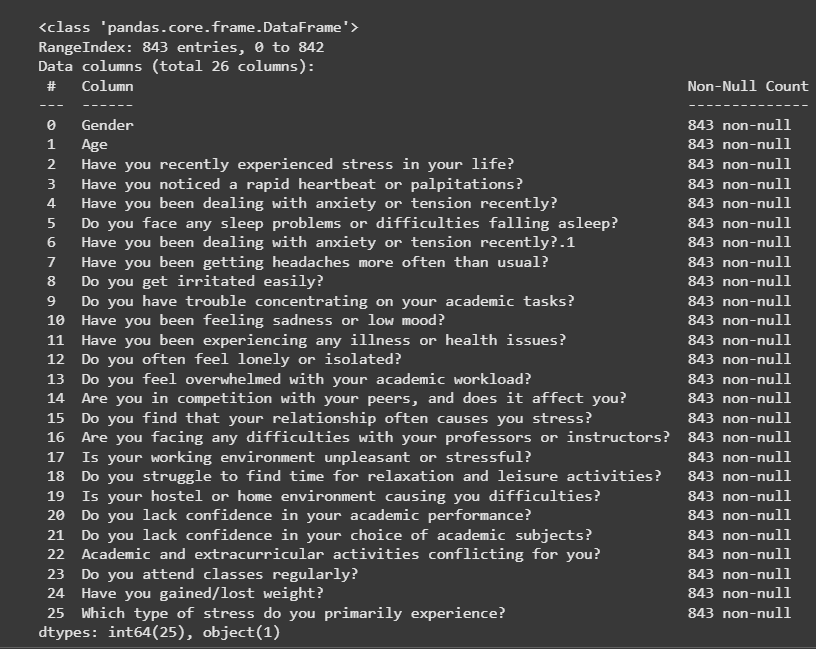
модели и научиться проводить базовый анализ важности признаков.

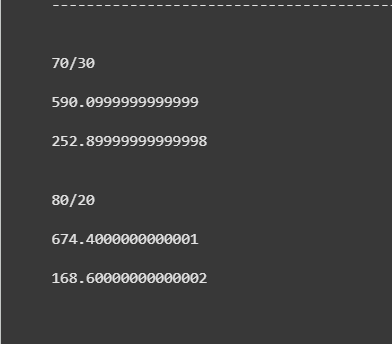
**Основная часть:**

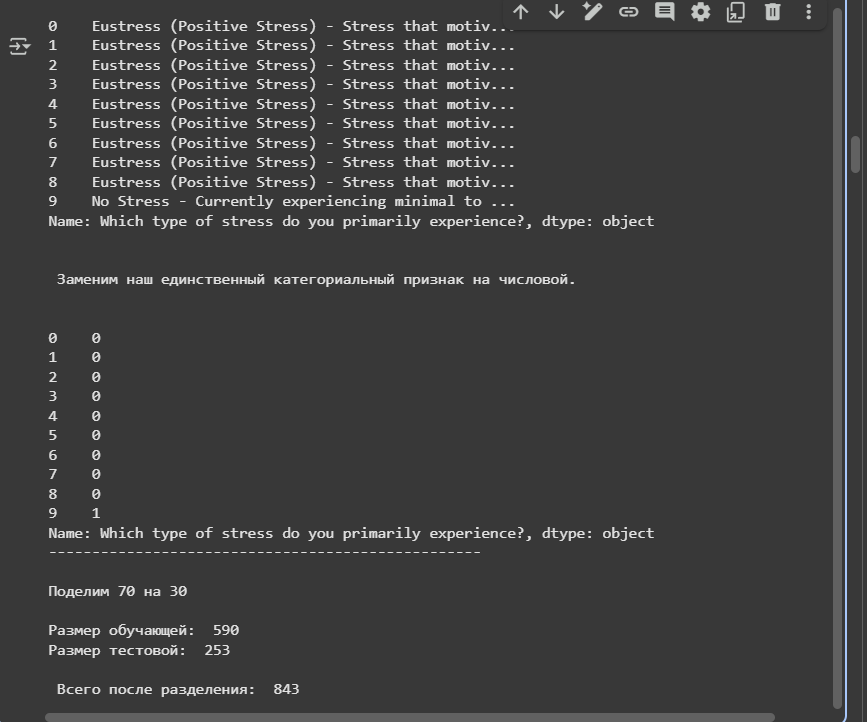
**1. Подготовка данных**

1.1. Выбираем для работы датасет из первой лабораторной работы, а конкретно файл Stress\_Dataset. За целевую переменную берём вопрос Have you recently experienced stress in your life?

1.2. Загружаем и выполняем предобработку.



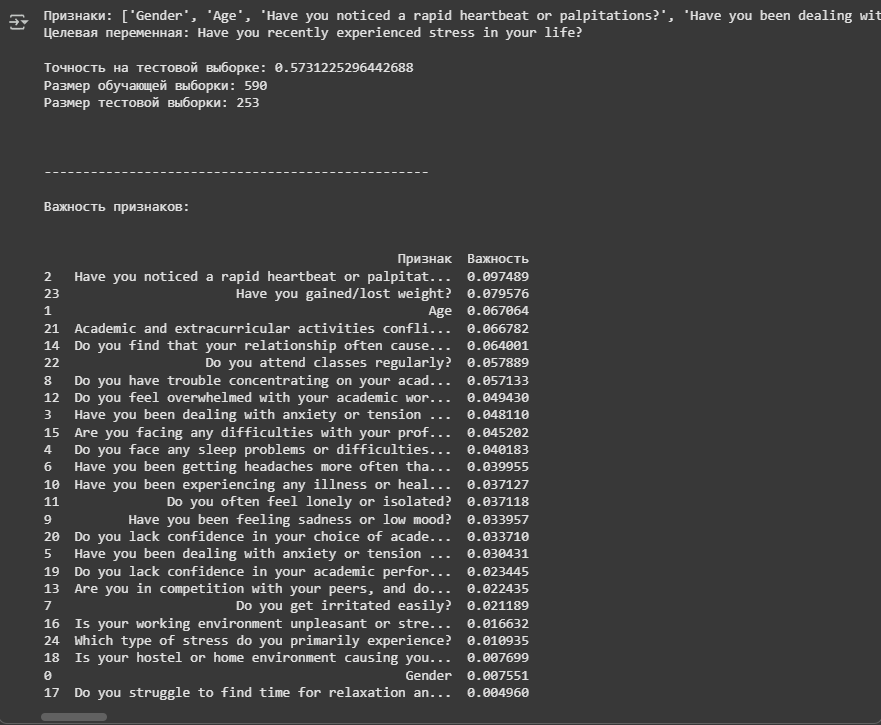




Итого: пропущенных значений нет, единственный категориальный признак в числовой, разделили выборку на обучающую и тестовую в соотношении 70 на 30.

**2. Базовое дерево**

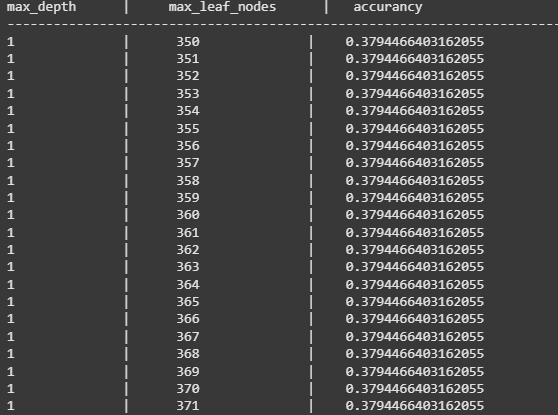


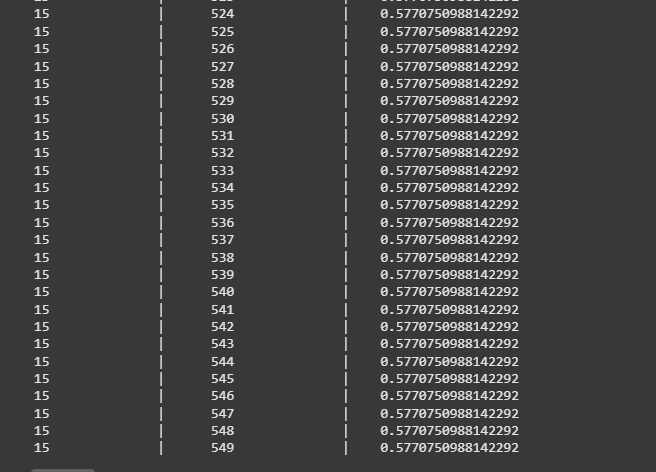


Итого: точность на тестовой выборке с параметрами по умолчанию примерно 57,3%; три наиболее важных признака: ускоренное сердцебиение, лишний вес и возраст.

**3. Подбор гиперпараметров**

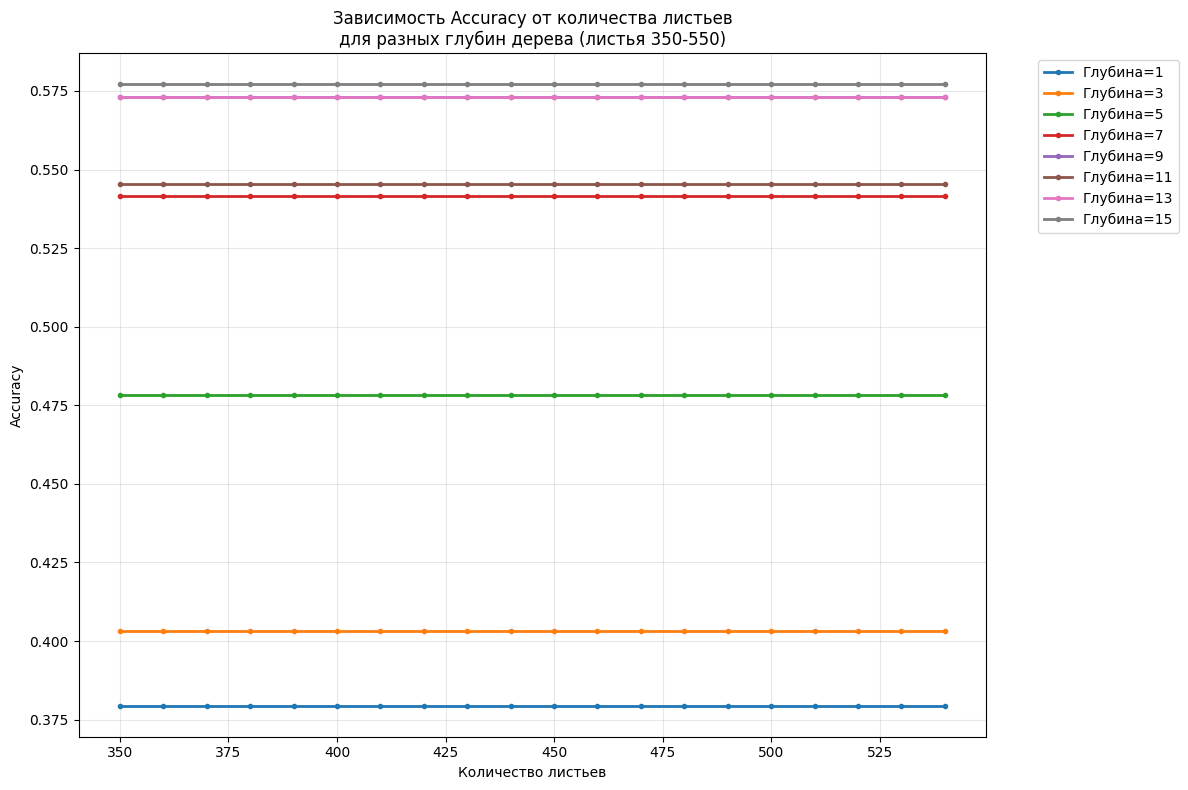


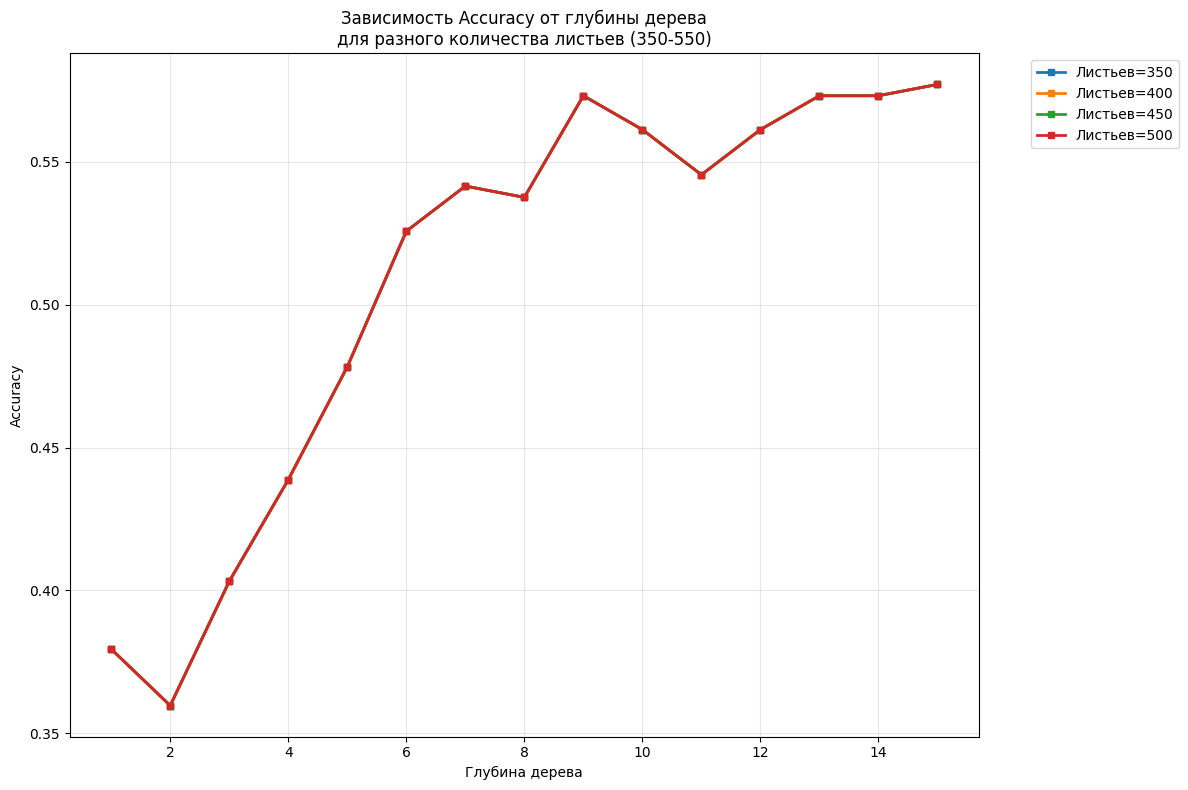




**4. Анализ результатов**

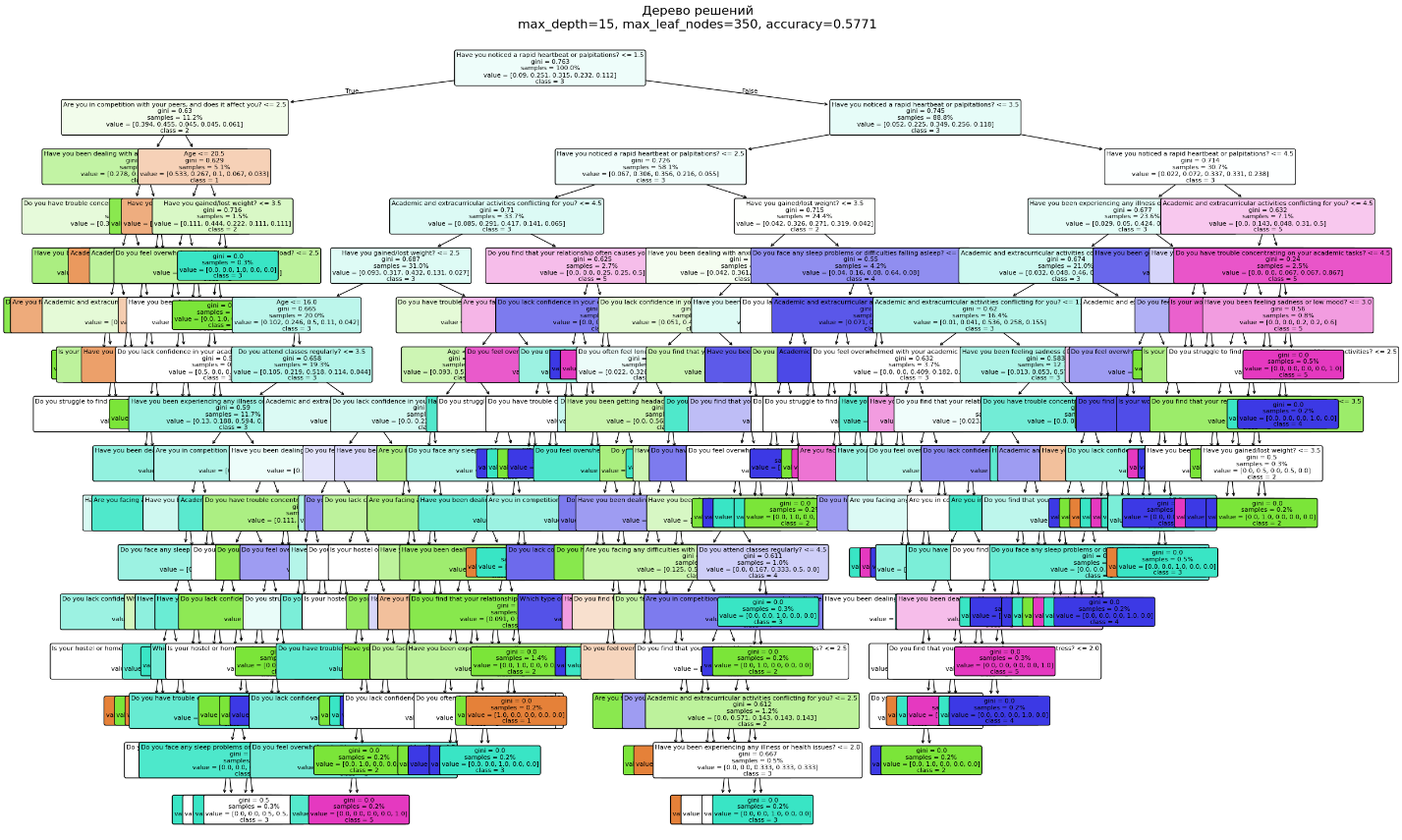
Максимальная точность 0.5770750988142292, то есть примерно 57.7%, на 0.04% больше, чем у базового дерева.





**5. Визуализация**

Полное дерево



Первые 4 уровня

