Федеральное государственное бюджетное образовательное учреждение высшего образования “Сибирский государственный университет телекоммуникаций и информатики”

Лабораторная работа по теме “Решающие деревья”

Выполнил:  
Студент группы ИП-712  
Коротеев Максим Сергеевич

Бахарев Максим Евгеньевич

Работу проверил:

ассистент кафедры ПМиК  
Морозова К.И.

# Задание

Реализовать классификатор на основе решающих деревьев.

# Процесс выполнения работы

1. Данные были разбиты на обучающий и тестовый наборы с соотношением 7:3
2. На обучающем наборе данных было создано решающее дерево.
3. На тестовом наборе данных было протестировано созданное решающее дерево.

* Точность на 10 разбиениях по умолчанию составила:  
  0.76, 0.77, 0.71, 0.76, 0.74, 0.76, 0.78, 0.72, 0.72, 0.72.
* Точность на 5 разбиениях с максимальной глубиной дерева равной 5 и максимальным количеством листьев 10:  
  0.79, 0.77, 0.76, 0.82, 0.81
* Точность на 5 разбиениях с максимальной глубиной дерева равной 10 и максимальным количеством листьев 5:  
  0.77, 0.84, 0.76, 0.8, 0.78

# Вывод

В лабораторной работе был реализован классификатор на основе решающих деревьев. При изменении глубины дерева и количества листьев можно наблюдать небольшое улучшение результатов. Классификатор на основе решающих деревьев показывает точность в данной задаче в районе 0.75.

# Листинг

|  |
| --- |
| #%% md  # Lab2  Обработка данных  #%%  import numpy as np  import sklearn as sk  from sklearn.impute import SimpleImputer  dataset = np.genfromtxt(**'heart\_data.csv'**, delimiter=**','**, skip\_header=True)  X = dataset[:,:-1]  y = dataset[:,-1]  X\_train, X\_test, y\_train, y\_test = sk.model\_selection.train\_test\_split(X, y, test\_size=0.3)  imp = SimpleImputer()  imp.fit(X\_train)  X\_train = imp.transform(X\_train)  #%% md  Обучение  #%%  from sklearn.tree import DecisionTreeClassifier  clf = DecisionTreeClassifier()  clf.fit(X\_train, y\_train)  #%% md  Проверка  #%%  X\_test = imp.transform(X\_test)  predicted = clf.predict(X\_test)  sk.metrics.accuracy\_score(y\_test, predicted) |