Министерство науки и высшего образования Российской Федерации

Федеральное государственное бюджетное образовательное учреждение

высшего образования

«Курский государственный университет»

Кафедра информационной безопасности

Направление подготовки **09.03.01 Информатика и вычислительная техника**

Направленность(профиль) **Автоматизированные системы обработки информации**

Факультет **Физики, математики, информатики**

Дисциплина **Системы искусственного интеллекта**

Форма обучения **очная**

**ОТЧЕТ**

по лабораторной работе №2

Выполнил: студент 2 курса Красников Я.И.  
Проверил: ассистент кафедры ИБ Юрьев И. А.

Курск – 2025

**Лабораторная работа № 2**

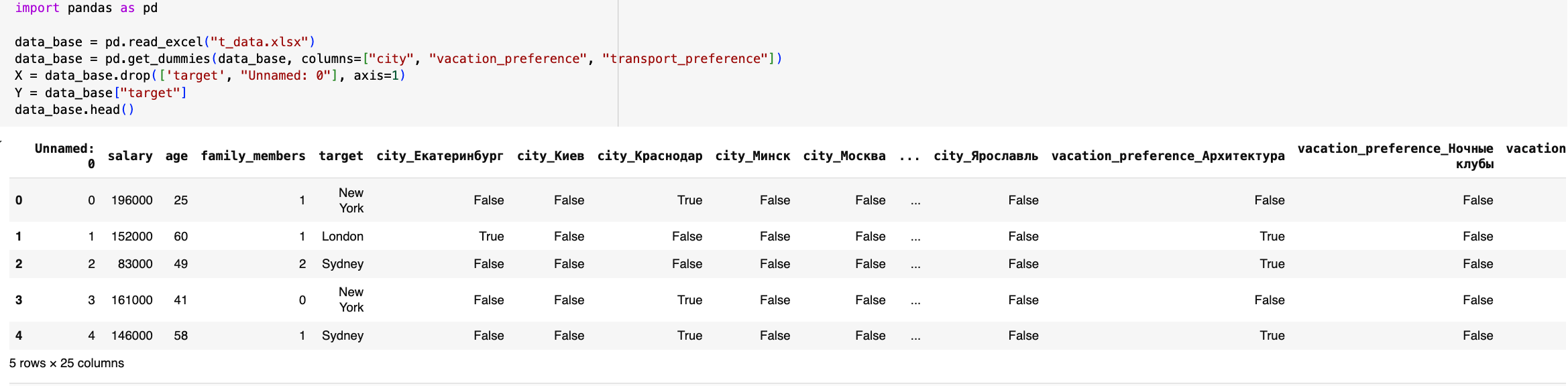
1. Категориальные и количественные переменные, дайте определения.

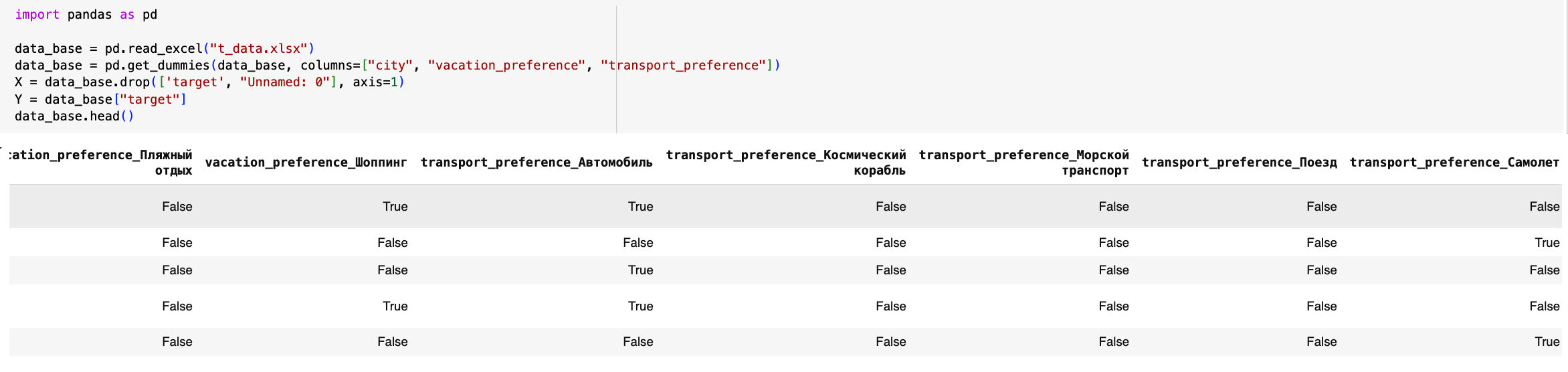
**Категориальная переменная** – это переменная с ограниченным числом уникальных значений или категорий. Обычно связана с неисчисляемыми атрибутами, такими как названия, имена людей, исходы событий и т.д.

Как правило, значения категориальных переменных являются строковыми. Но иногда могут использоваться и числа, если некоторое наименование кодируется числовым значением.  
Например, вместо категорий «Провальный», «Средний», «Успешный» в модели оценки успешности продукта можно использовать значения 0, 1, 2.

**Количественная переменная** – это переменная, которая может принимать любые числовые значения в некотором диапазоне. Количественные переменные можно вычислить, а также можно агрегировать. Вычислить сумму или среднее, к примеру.

1. Произведите предобработку данных путем обработки категориальных признаков



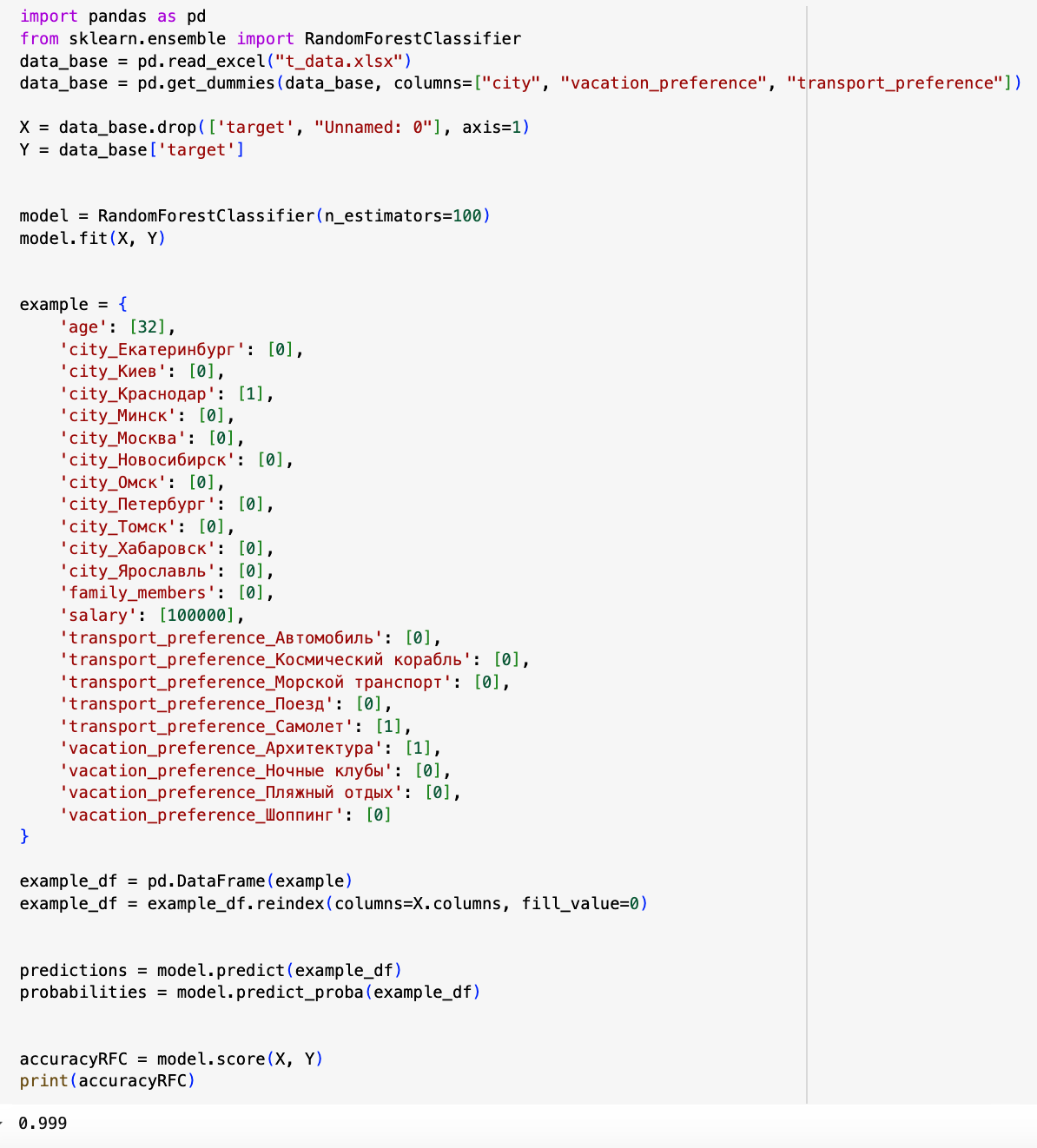


1. Что такое One-hot encoding?

One-hot encoding — это процесс преобразования категориальных переменных в двоичный формат. Он создаёт новые столбцы для каждой категории, где 1 означает, что категория присутствует, а 0 означает, что её нет.

1. Используя исходный код, настройте гиперпараметры модели.

Гиперпараметры модели — это настройки, которые определяют общую структуру и способ обучения модели. Они устанавливаются до начала процесса обучения и не изменяются в нём.



1. Попробуйте, используя сайт https://scikit-learn.org/ выбрать 3 другие модели, и применить их для работы с данными.

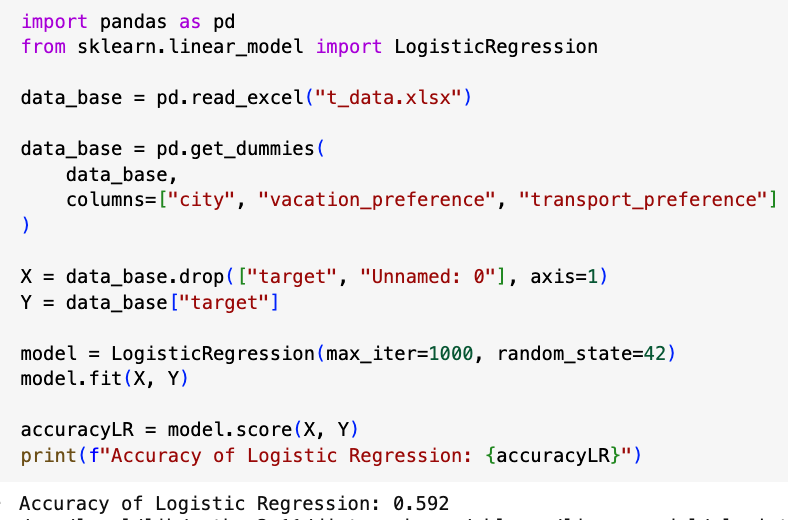
SVC (Support Vector Classifier) является частью семейства методов опорных векторов (SVM, Support Vector Machines). Основная идея SVC заключается в нахождении гиперплоскости, которая разделяет данные на классы с максимальным отступом (маржином) между ними.



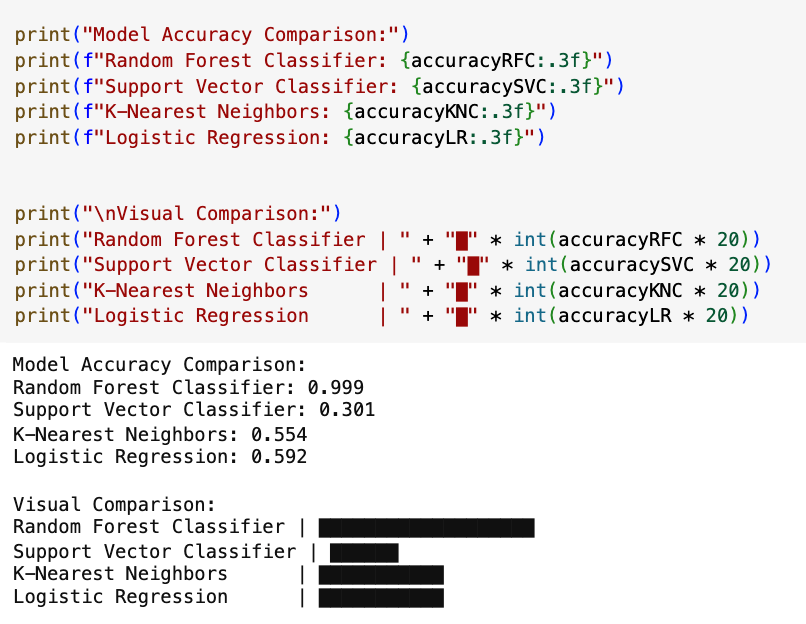
KNeighbors (K-Nearest Neighbors, KNN) основан на идее, что объекты, находящиеся ближе друг к другу в пространстве признаков, с большей вероятностью принадлежат к одному и тому же классу.



LogisticRegression (Логистическая регрессия) — это статистический метод, используемый для задач бинарной классификации, то есть когда нужно предсказать одну из двух возможных категорий. Он моделирует вероятность принадлежности объекта к одному из классов.



1. Сравните полученные точности.



Ссылка на github:

Ссылка на colab: