|  |
| --- |
|  |
| МИНОБРНаУКИ РОССИИ |
| Федеральное государственное бюджетное образовательное учреждение  высшего образования  **«МИРЭА – Российский технологический университет»**  **РТУ МИРЭА** |

Институт комплексной безопасности и специального приборостроения

**Отчет** по лабораторной работе №1

**по дисциплине:** «Анализ защищенности систем искусственного интеллекта»

**Выполнил:**

Студент группы ББМО-01-22

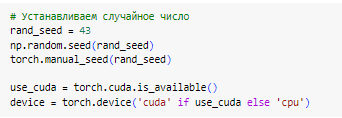
ФИО: Карпова Н.Ю.

**Москва 2023**

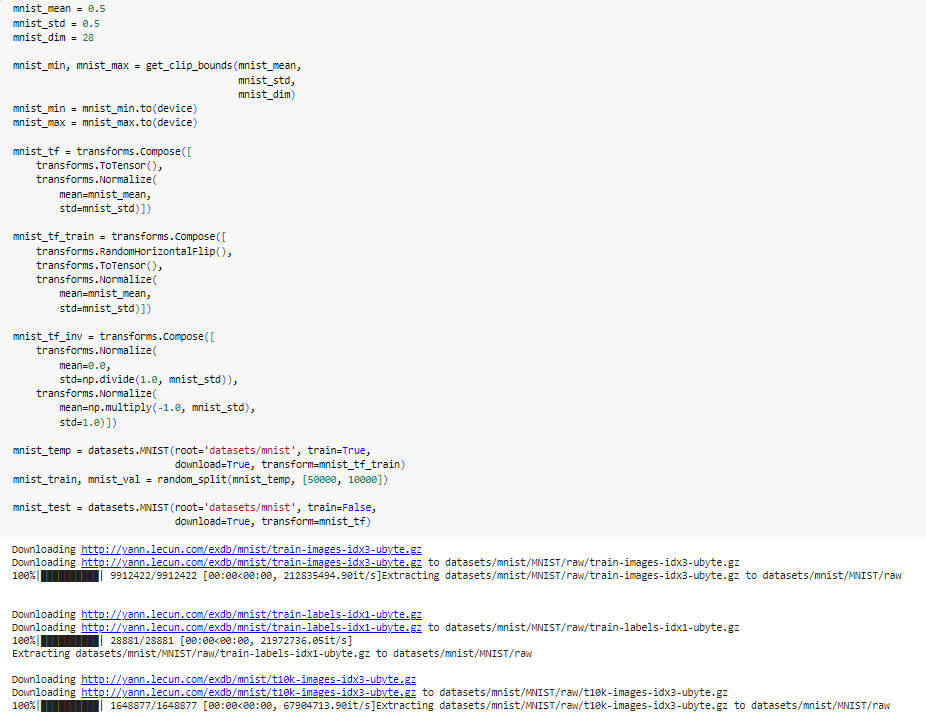
1. Клонируем репозиторий и загружаем нужные библиотеки.



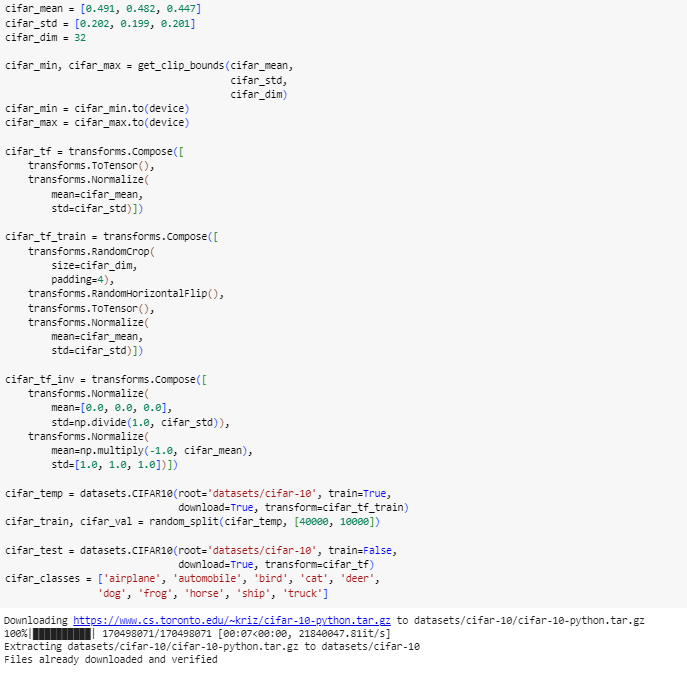
1. Устанавливаем случайное число и выбираем устройство выполнения.



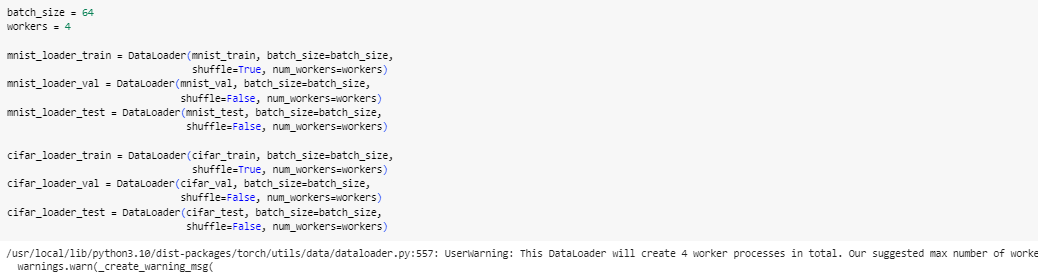
1. Загружаем dataset «MNIST» и отредактируем его.



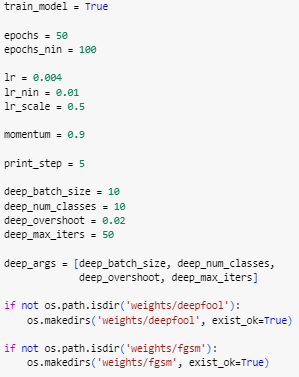
1. Загружаем dataset «CIFAR-10» и отредактируем его.



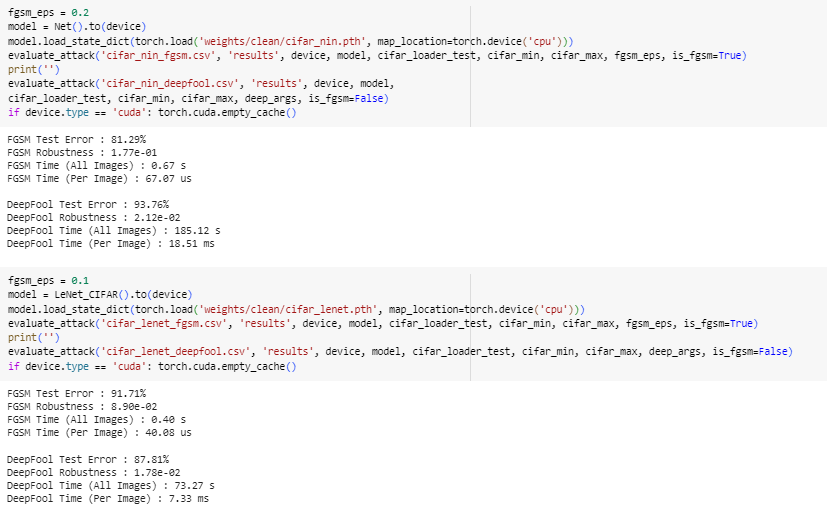
1. Отредактируем гиперпараметры.



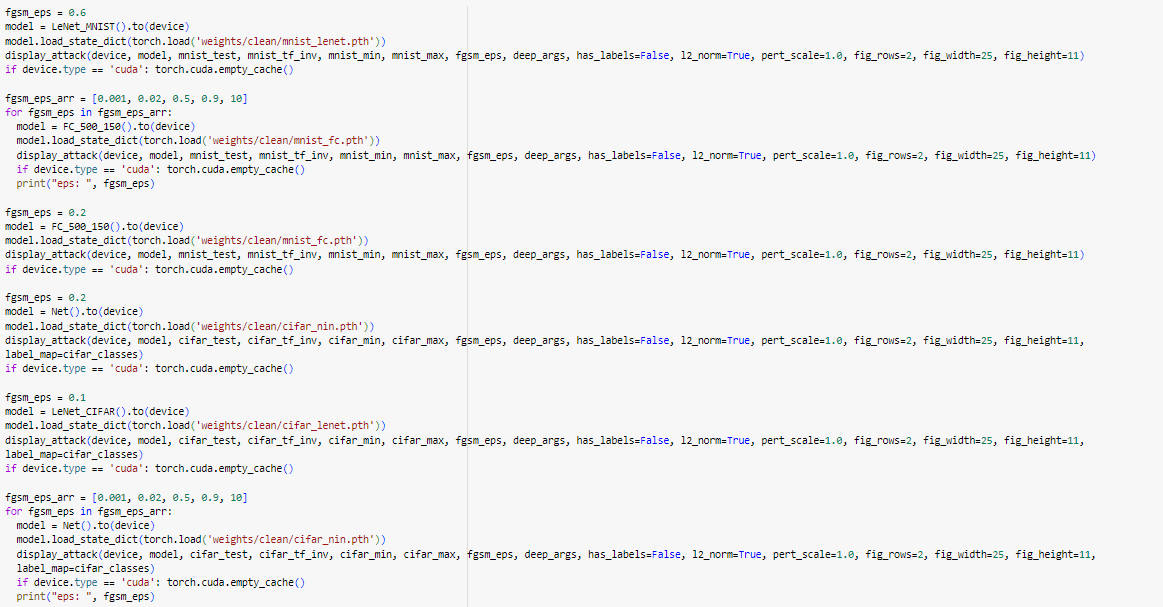
1. Зададим параметры на модель.



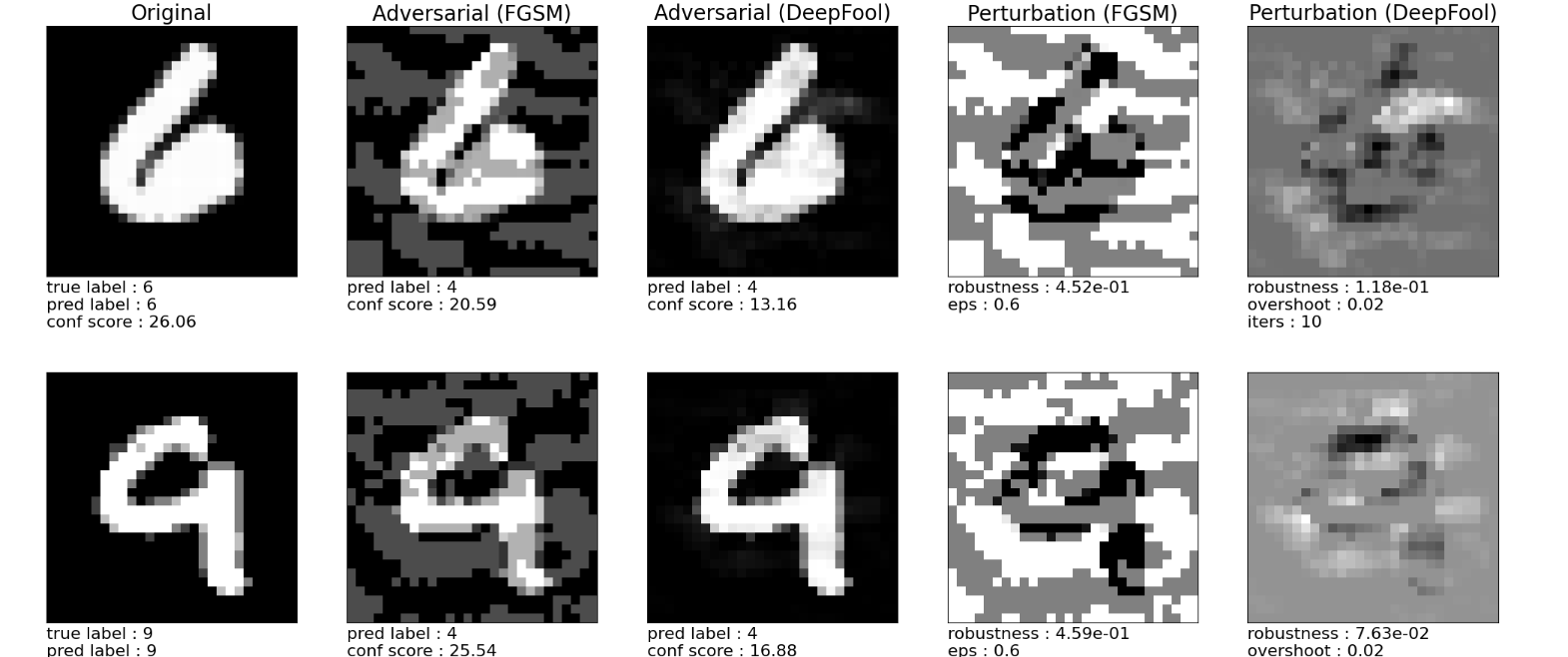
1. Загрузим и оценим стойкость модели.

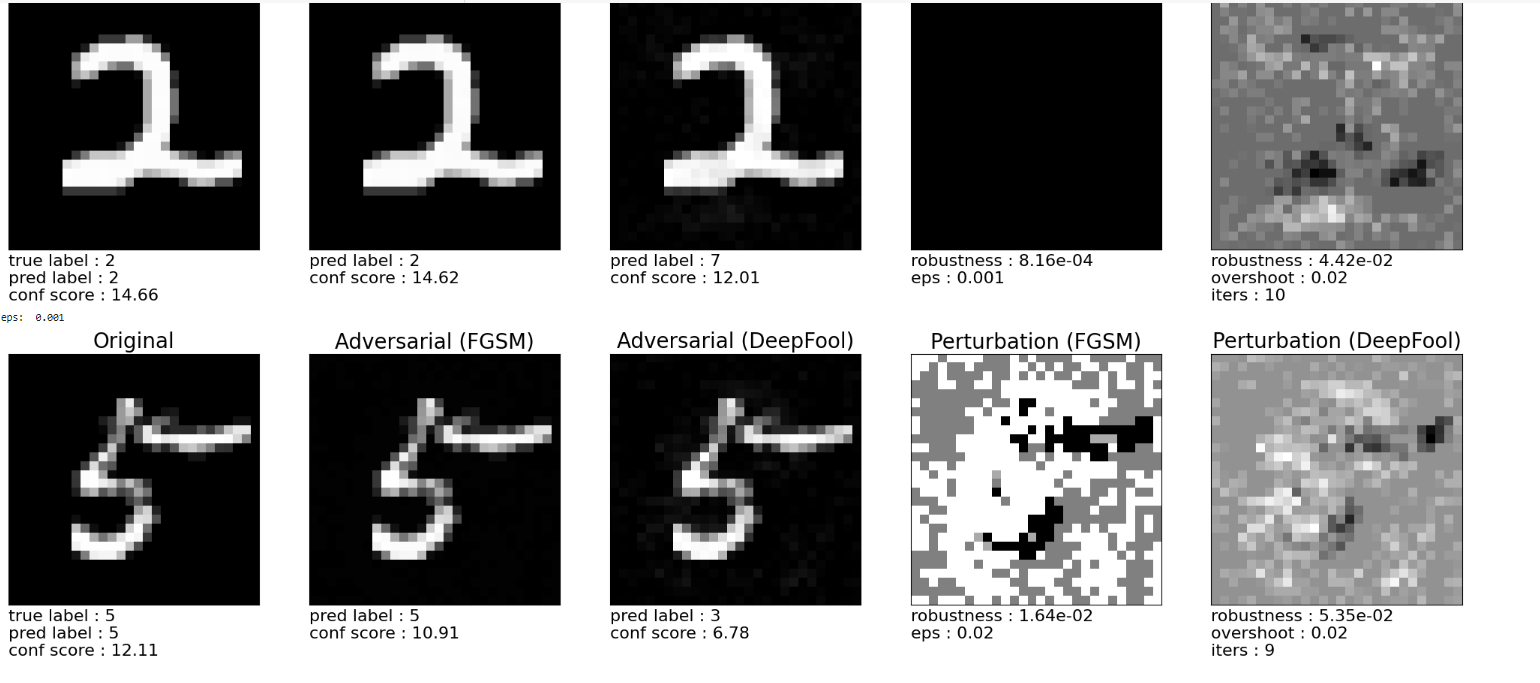


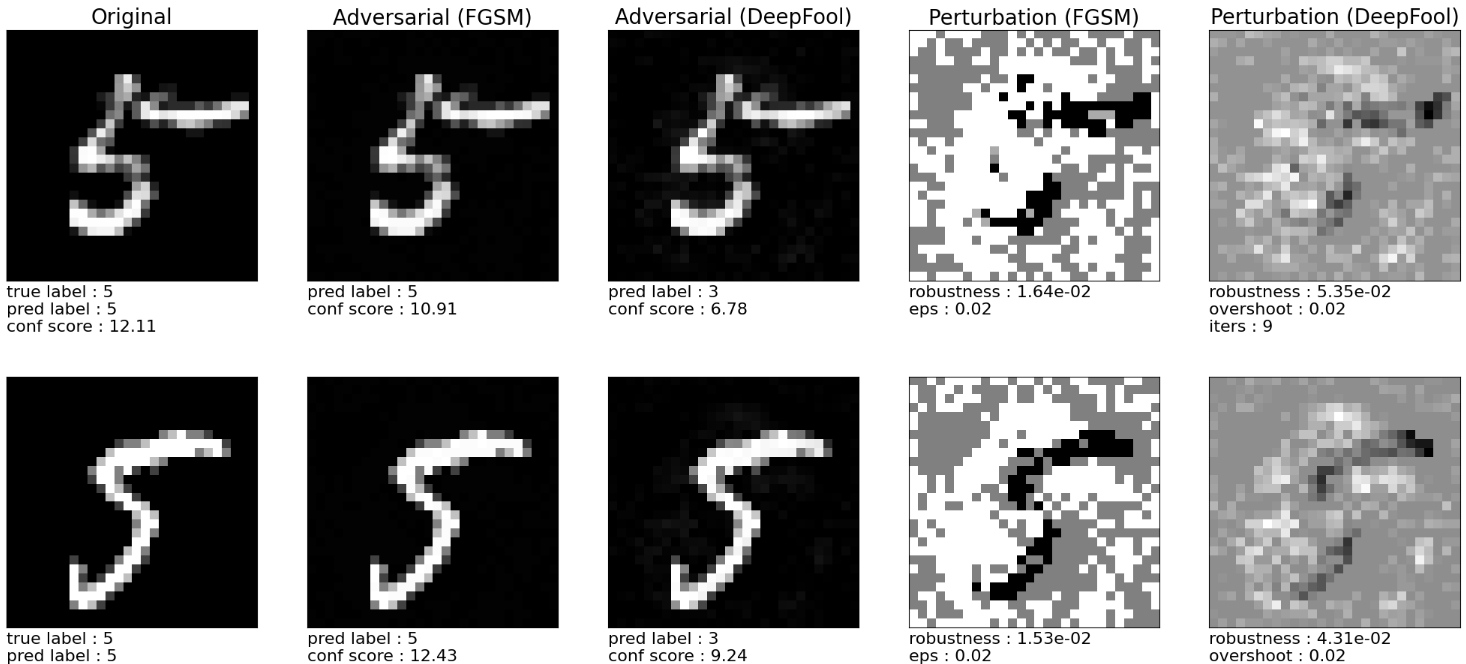
1. Выполним оценку атакующих примеров для сетей.

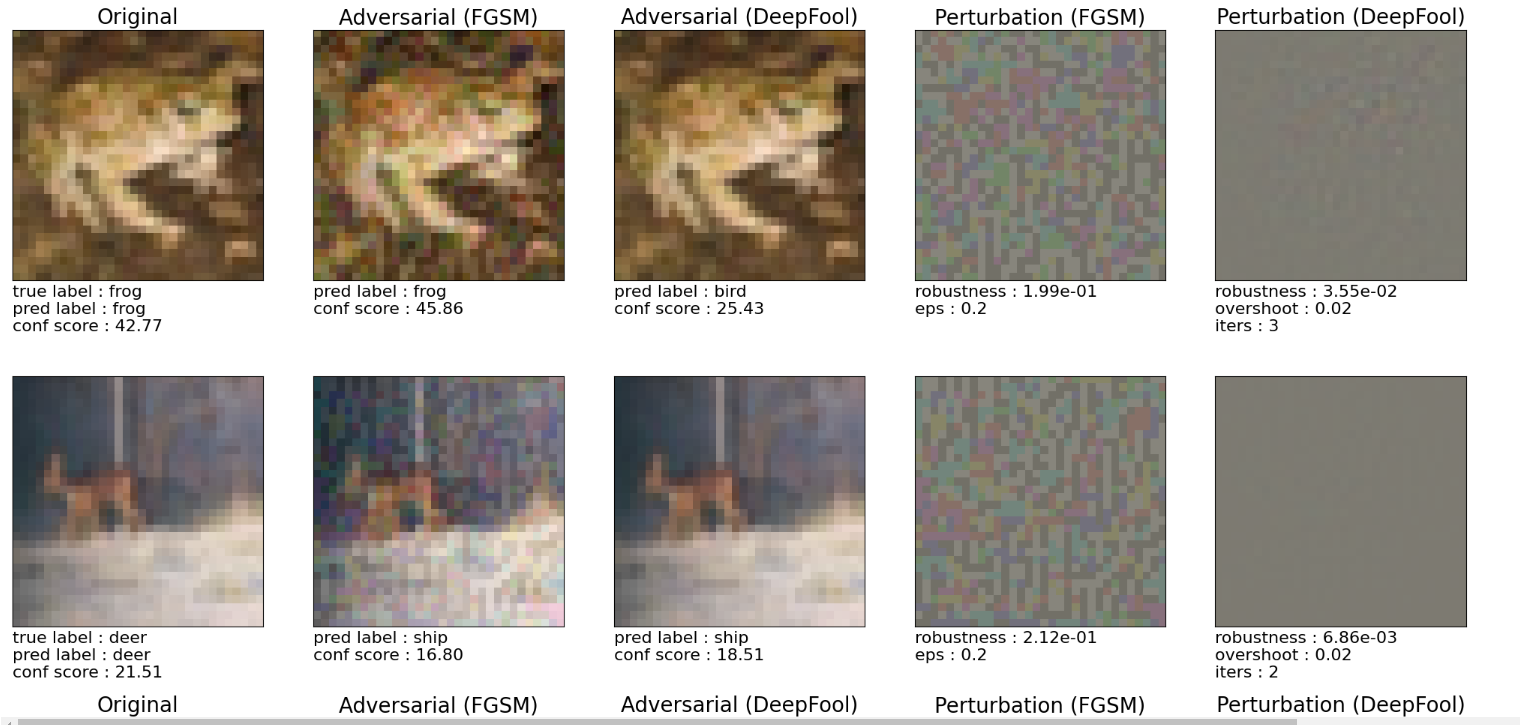


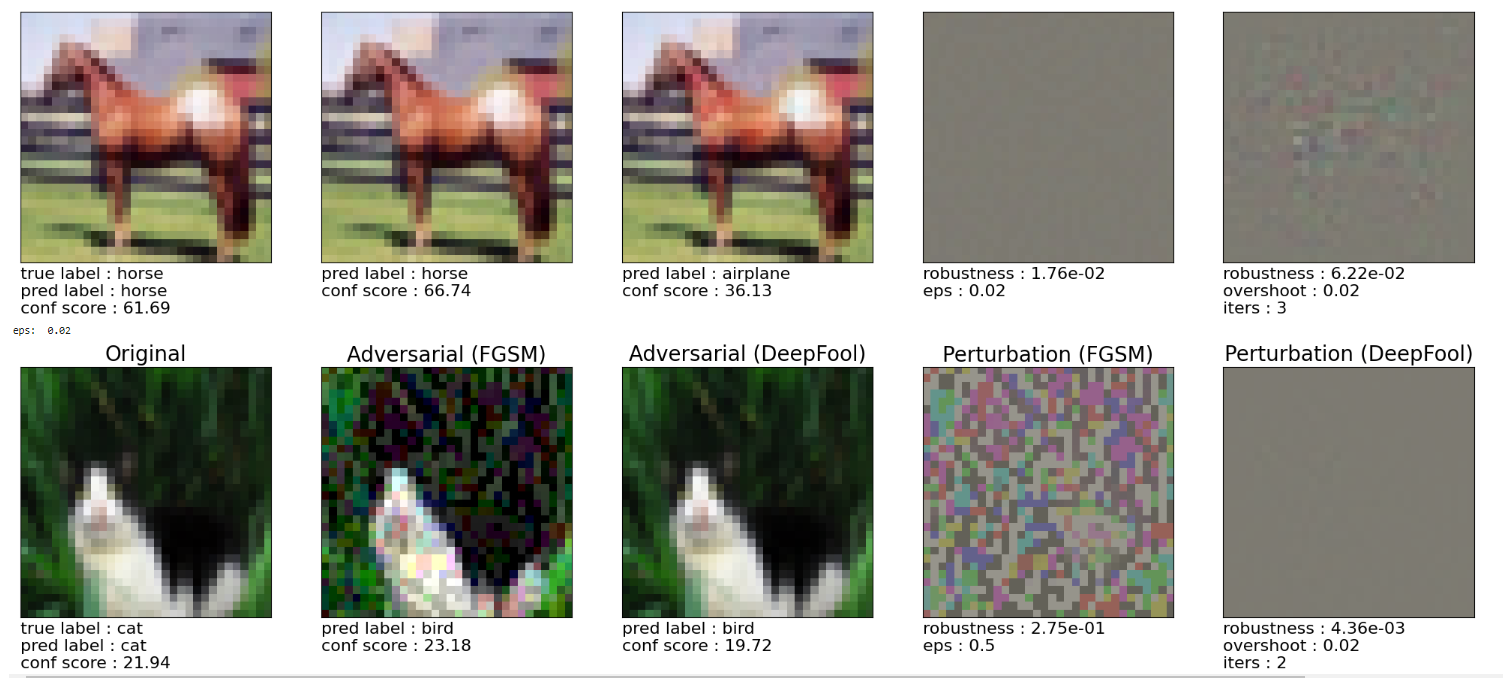
1. Оценим результат атаки.

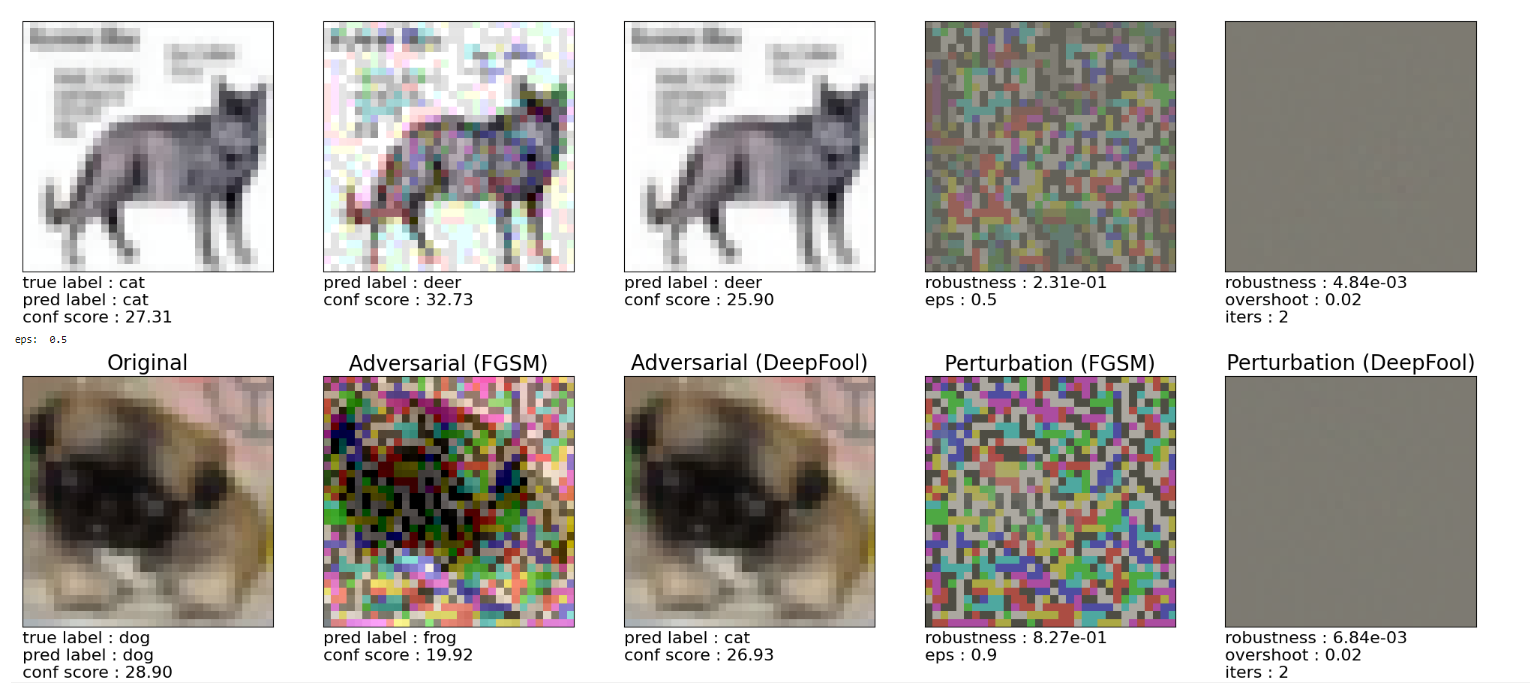


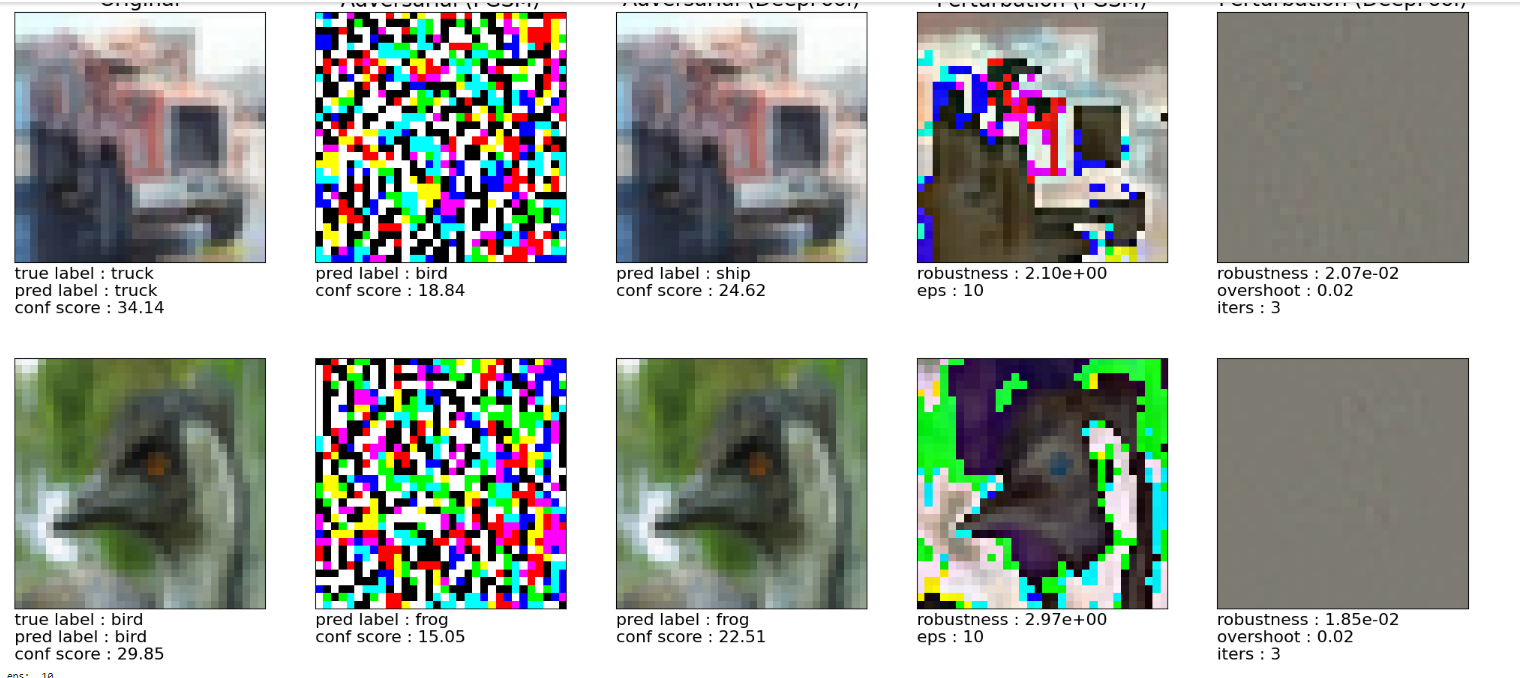












**Выводы**

Из результатов выполнения лабораторной работы 1 следует, что малые значения параметра fgsm\_eps продемонстрировали сохранение устойчивости сетей к атакам, при этом уровень ошибок классификации оставался на низком уровне. Однако, при увеличении значения fgsm\_eps сети становились более уязвимыми к атакам и начинали допускать больше ошибок в классификации. Эти наблюдения были особенно выражены на сети FC LeNet, используемой для датасета MNIST, а также на сети NiN LeNet, применяемой для датасета CIFAR.

Эти результаты приводят к важному выводу о влиянии параметра fgsm\_eps на стойкость сетей к атакам. В частности, они подчеркивают необходимость баланса между уровнем защиты и точностью классификации при выборе значения этого параметра. При выборе слишком большого значения fgsm\_eps сети могут стать слишком уязвимыми к атакам, что может привести к неприемлемым ошибкам в классификации. С другой стороны, слишком маленькие значения fgsm\_eps могут не обеспечить достаточного уровня защиты от атак.

Таким образом, необходимо тщательно настраивать параметр fgsm\_eps, учитывая специфику задачи и требования к безопасности модели. Дополнительные исследования и эксперименты могут быть проведены для более глубокого понимания влияния этого параметра на стойкость сетей к атакам и точность классификации.