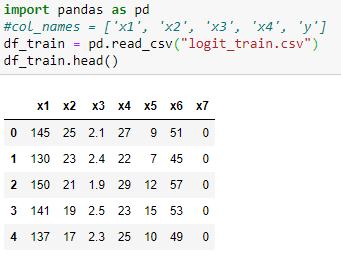
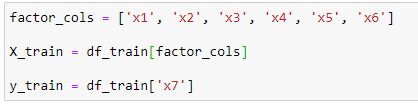
**Логистическая регрессия.**

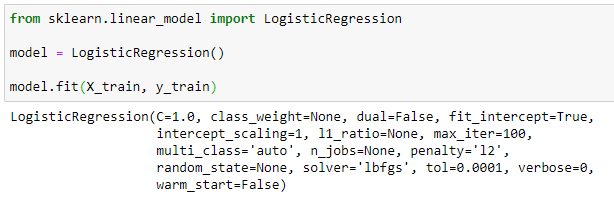
Считываем данные:

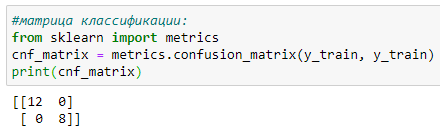


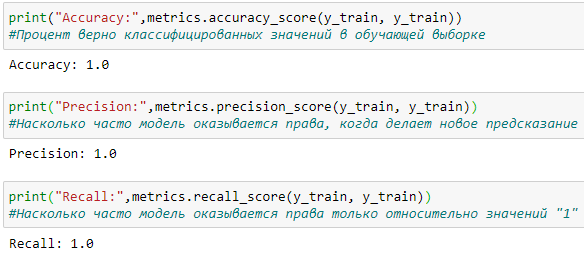
Разделяем данные на целевую переменную (х7) и признаки (х1-х6)



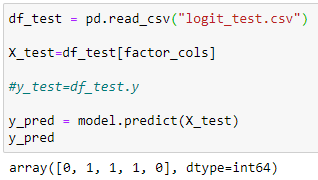
Создаём модель:



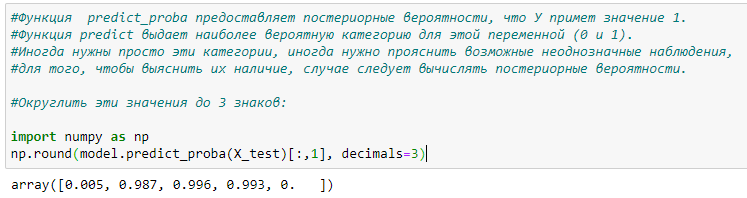


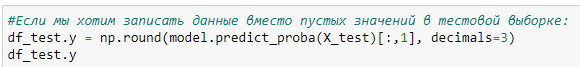


Считываем тестовые данные, записываем их в «признаки». Затем предсказываем «целевую» переменную на основе обученной модели:



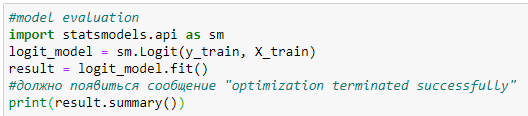
Собственно, это и есть отнесение новых инвестиционных проектов к группам. Первый и пятый проект – «успешны», остальные – «неуспешны».







При попытке запустить код:



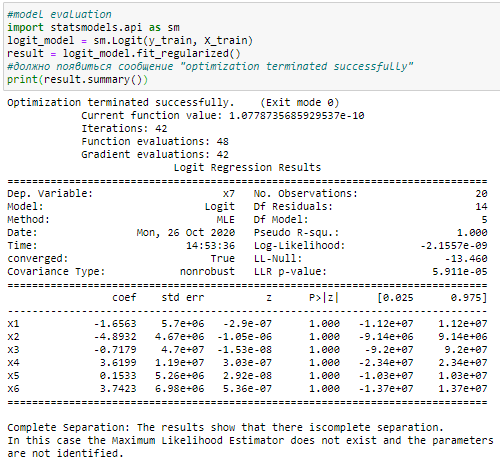
Возникает ошибка:



Которая говорит о том, что модель обнаружила полное (или почти полное) разделение по одному из признаков. Это происходит, когда все (или почти все) значения какого-либо признака (или комбинации) связаны только с одним из двоичных значений целевой переменной. В этом случае решение не может быть найдено для коэффициентов.

То есть, другими словами, целевая переменная полностью зависит от одного из признаков (От x3 - коэффициента покрытия ссудной задолженности, если судить по данным).

Для того, чтобы эта ошибка не возникала, нужно большее количество данных, либо можно использовать регуляризацию:



Однако, в конце всё равно выводится сообщение, что обнаружено полное разделение и метод максимального правдоподобия (На основе которого производится расчёт коэффициентов) не может дать результатов.

Строим график кривой регрессии:

