3.1

Чисто теоретически модель с accuracy 95% может со 100% точностью определять мошеннические заявки, но в 7% считать, что человек мошенник, когда это не так. Если для нашего банка 7% проверок добропорядочных клиентов выгоднее чем дать деньги 2% мошенников , тогда наша модель полезна.

Также наша модель может быть абсолютно бесполезна и вредна т.к Accuracy это доля верных предсказаний, а у нас одного предсказываемого класса 98%, а другого 2%, поэтому модель которая просто предсказывала бы всегда, что заявка не мошенническая имела бы accuracy ~=98%, поэтому в худшем случае наша модель никогда не определяет мошенников так ещё и неверно определяет мошенниками добропорядочных клиентов в 3% случаев, в таком случае мы даже понесём большие убытки, чем просто выдавали бы кредиты всем.

3.2

Точные предсказания лежат в квадратах (мошенник, мошенник) и (не мошенник, не мошенник)

Accuracy = (77 + 9487) / (77 + 127 + 374 + 9487) = 0.9502235469448584

Кол-во мошенников это просто сумма мошенников в результате проверки

Доля мошенников в выборке = (77 + 127) / (374 + 9487 + 77 + 127)= 0.020268256333830104

3.3

Сколько мы заработаем (сэкономим), если будем использовать модель?

Когда мы неправильно определяем мошенника мы теряем 500 рублей

Когда мы правильно определяем мошенника мы экономим 400 рублей

Когда мы правильно определяем не мошенника мы экономим 100 рублей

Когда мы неправильно определяем не мошенника мы теряем 100 рублей

Тогда заработок(экономия) будет равен: (400\*кол-во правильно определенных мошенников)+(100\*кол-во правильно определенных не мошенников ) - (500\*кол-во неправильно определенных мошенников) – (100\*кол-во неправильно определенных не мошенников)

В прошлом месяце мы бы сэкономили (400\*77) +(100\*9487) – (500\*127) – (374\*100) = 878600

Сколько мы заработаем (сэкономим), если вместо внедрения модели перестанем проводить проверки вообще?

Когда мы не проводим проверку мы экономим 100 рублей

Когда приходит мошенник теряем 500 рублей

Заработаем или сэкономим (100\*кол-во клиентов) – (500\*кол-во мошенников), в прошлом месяце мы бы сэкономили (9487 + 127 + 77 + 374) \* 100 - 500 \* (77 + 127) = 904500

3.4

Наша экономия в случае отключения всех проверок зависит от доли мошенников, если мы отключаем полностью все проверки, то вероятно мошенники это поймут и начнут чаще пользоваться услугами для мошеннических целей, плюс об этом узнают другие, тогда доля мошенников сильно вырастет и как следствие мы понесём убытки.