



Medicina para todos

Поиск мошеннических
транзакций в добровольном
медицинском страховании

Система медицинского страхования в Кабо-Верде

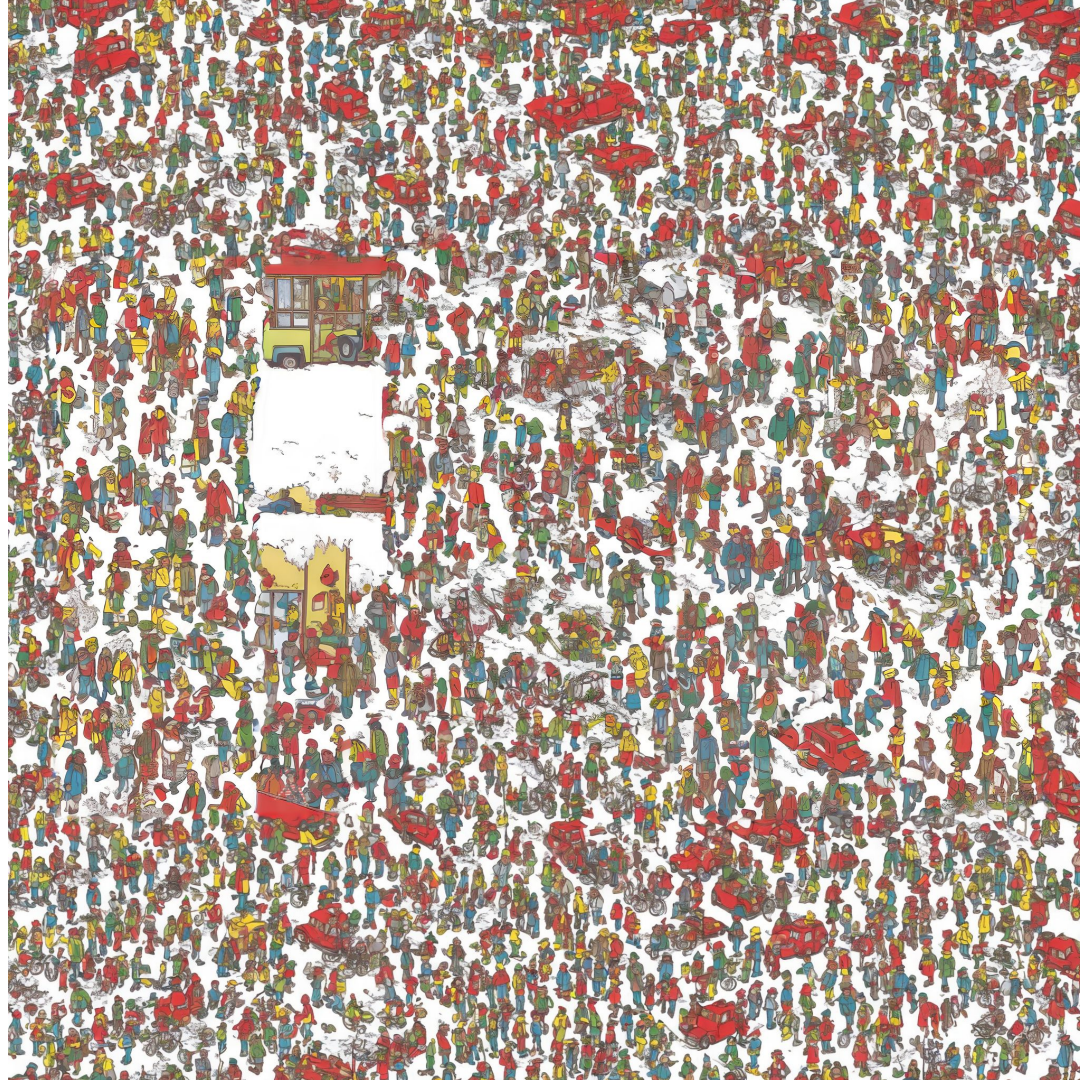


Старение населения, популярность вредной пищи, сидячий образ жизни

Пандемии, эпидемии

Значимое количество мошеннических транзакций со стороны контрагентов





Find Waldo

Как выглядит поиск
мошеннической
транзакции

Почему это сложно и важно?

Обман распространен!

По статистике около 20% счетов, выставляемых локальными поликлиниками по программам добровольного медицинского страхования, содержат услуги, которые либо не нужны пациентам, либо физически им не оказаны.

Дороговизна экспертизы

Проверка корректности счетов и назначений может осуществляться только сотрудниками с медицинским образованием. К тому же невозможно проверить 100% вручную, проверка проводится выборочно



Обман маскируется

Разнообразие мошеннических схем, клиники адаптируют и изобретают новые способы мошенничества с выставлением счетов за медицинские услуги

№	Type of fraud	Definition	Beneficiary
1	Phantom billing	Providers submit invoices for services that were never performed	Medical Services Providers
2	Claims Duplication	Providers charge twice for the same service.	
3	Claims Splitting	Providers split one medical episode into two claims to avoid detection.	
4	Upcoding	Providers charge for more expensive services than were actually performed.	
5	Unbundling	Providers charge for individual services that are part of a larger service.	
6	Misdiagnosing	Providers submit incorrect ICD codes* to charge for more severe diseases.	Insured person and a doctor
7	Identity theft	Individuals use someone else's ID to receive healthcare services.	
8	Kickbacks	Providers offer payment or compensation to patients to encourage them to come to their office.	
9	Doctor shopping	Patients visit multiple doctors to obtain <u>drugs</u> prescriptions.	Insured person
10	Approver & Provider collusion	Providers pay insurers' employees to approve more requests.	Employees of an insurer and provider
11	Claim Handler & Provider collusion	Providers pay claim handlers to approve more claims.	

Medicina para todos

Лидер локального рынка страхования медицинских услуг.

Компания испытывает трудности в связи со значимой долей мошеннических операций, которые являются следствием недобросовестности как некоторых сотрудников компании, так и контрагентов: пациентов, врачей, сотрудников лечебных учреждений.

14 000 000

Пациентов по всей стране

8 000

Партнерских ЛПУ

6 290

Эскудо – средняя сумма выплат за одного застрахованного

46 500

Счетов в обработке ежемесячно

Задача

Снижение затрат на обеспечение проверки страховых случаев, которая в настоящее время проводится вручную с использованием труда высокооплачиваемых специалистов.

В результате внедрения модели планируется первичную проверку осуществлять автоматически и подключать специалистов только на те случаи, которые признаны первичной проверкой мошенническими или “пограничными”, что позволит сократить затраты на специалистов пропорционально доли мошеннических операций в общем объеме, то есть ориентировочно в 10 раз.

Описание данных

- BenetID - идентификатор пациента в системе

- DOB - дата рождения пациента
- DOD - дата смерти пациента (если пациент скончался на момент составления датасета, в случае, если пациент жив - N/A)
- Gender - пол пациента
- Race - раса (от 1 до 5)
- State, County - штат, округ
- IPAnnualReimbursementAmt, OPAnnualReimbursementAmt - суммы запрошенных страховых возмещений для случаев, требующих госпитализации и для случаев амбулаторного лечения
- IPAnnualDeductibleAmt, OPAnnualDeductibleAmt - суммы, которые пациент оплатил самостоятельно, соответственно при госпитализации и при амбулаторном лечении

Данные о пациентах

- ClaimID - идентификатор заявки (страхового случая)
- ClaimStartDt/ ClaimEndDt - дата начала/окончания страхового случая
- Provider - идентификатор сотрудника страховой компании, который принимал решение по данному случаю
- InscClaimAmtReimbursed - сумма возмещения по данному случаю
- DeductibleAmtPaid - сумма, уплаченная пациентом
- AttendingPhysician, OperatingPhysician, OtherPhysician - идентификаторы врачей, участвовавших на разных стадиях лечения
- AdmissionDt/ DischargeDt - дата поступления/выписки пациента в стационар

Данные о госпитализациях

- ClaimID - идентификатор заявки (страхового случая)
- ClaimStartDt - дата начала страхового случая
- ClaimEndDt - дата окончания страхового случая
- Provider - идентификатор сотрудника страховой компании, который принимал решение по данному случаю
- InscClaimAmtReimbursed - сумма возмещения по данному случаю
- DeductibleAmtPaid - сумма, уплаченная пациентом
- AttendingPhysician, OperatingPhysician, OtherPhysician - идентификаторы врачей, участвовавших на разных стадиях лечения

Данные об амбулаторном лечении

- Provider - идентификатор сотрудника страховой компании
- PotentialFraud - категориальный идентификатор фрода (Yes, No)

Данные проверенных ранее транзакций

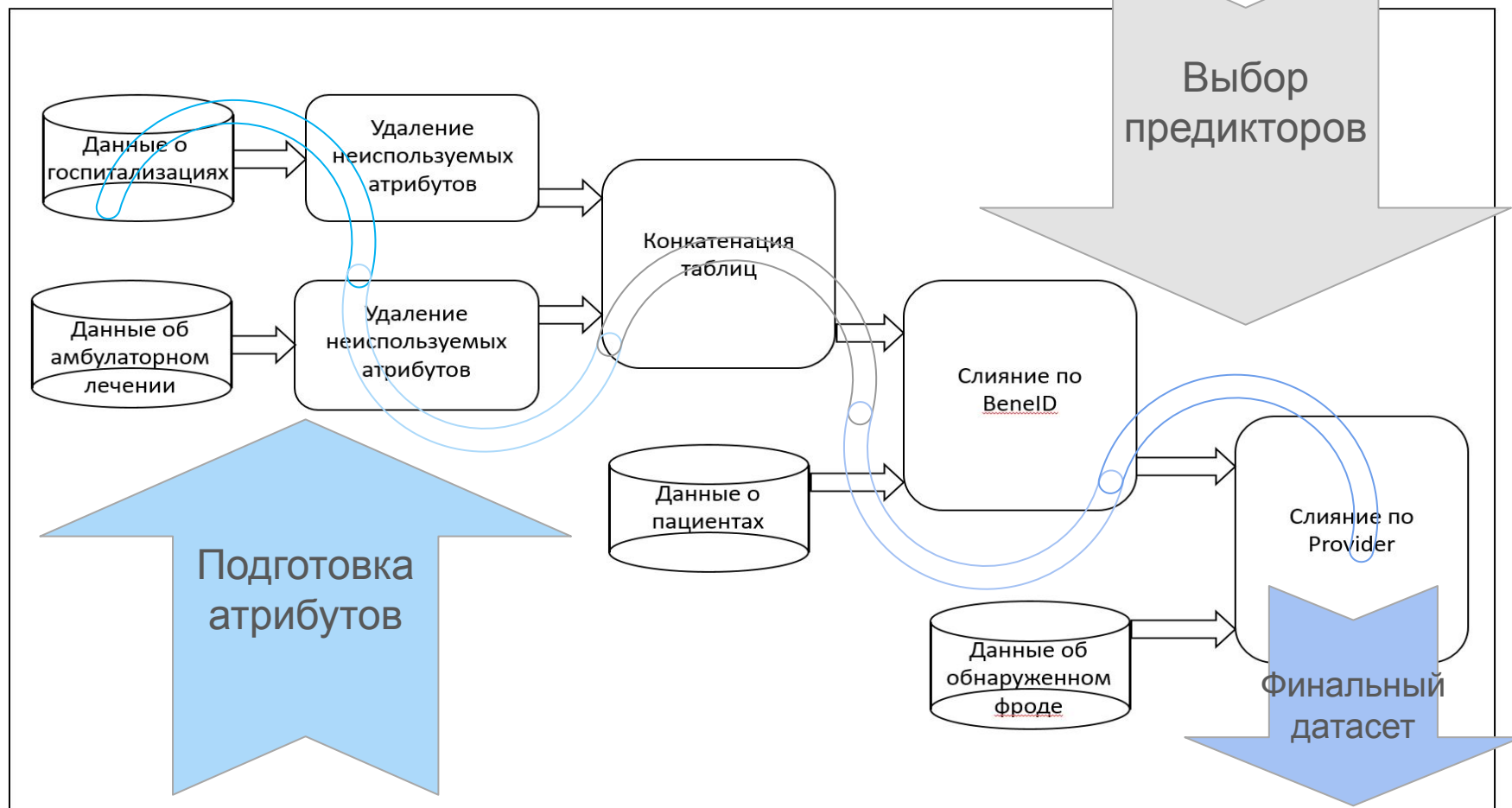
Внешние данные

DIAGNOSIS	ClaimID - Count
NEUROSES EXCEPT DEPRESSIVE	111
OTHER RESPIRATORY SYSTEM O.R. PROCEDURES WITH MCC	107
PLEURAL EFFUSION WITH MCC	102
CHRONIC OBSTRUCTIVE PULMONARY DISEASE WITHOUT CC/MCC	98
ORGANIC DISTURBANCES AND INTELLECTUAL DISABILITY	97
PLEURAL EFFUSION WITHOUT CC/MCC	97
DISORDERS OF PERSONALITY AND IMPULSE CONTROL	96
O.R. PROCEDURES WITH DIAGNOSES OF OTHER CONTACT WITH HEALTH SERVICES WITH MCC	96
BRONCHITIS AND ASTHMA WITH CC/MCC	95
REHABILITATION WITH CC/MCC	95
AFTERCARE WITH CC/MCC	94
OTHER RESPIRATORY SYSTEM O.R. PROCEDURES WITHOUT CC/MCC	94
RESPIRATORY SIGNS AND SYMPTOMS	94
PSYCHOSES	92
AFTERCARE WITHOUT CC/MCC	90
O.R. PROCEDURES WITH PRINCIPAL DIAGNOSIS OF MENTAL ILLNESS	90
OTHER MENTAL DISORDER DIAGNOSES	90
INTERSTITIAL LUNG DISEASE WITH MCC	89
INTERSTITIAL LUNG DISEASE WITHOUT CC/MCC	89
MAJOR CHEST TRAUMA WITH MCC	89

Топ-20 диагнозов, по которым было обнаружено максимальное количество фрода.

Коды диагнозов из датасета сопоставили с классификатором диагнозов

Подготовка данных



Моделирование

Decision Tree Orange Data Mining

- высокая интерпретируемость
- недостаточная точность
- ограниченность предикторов

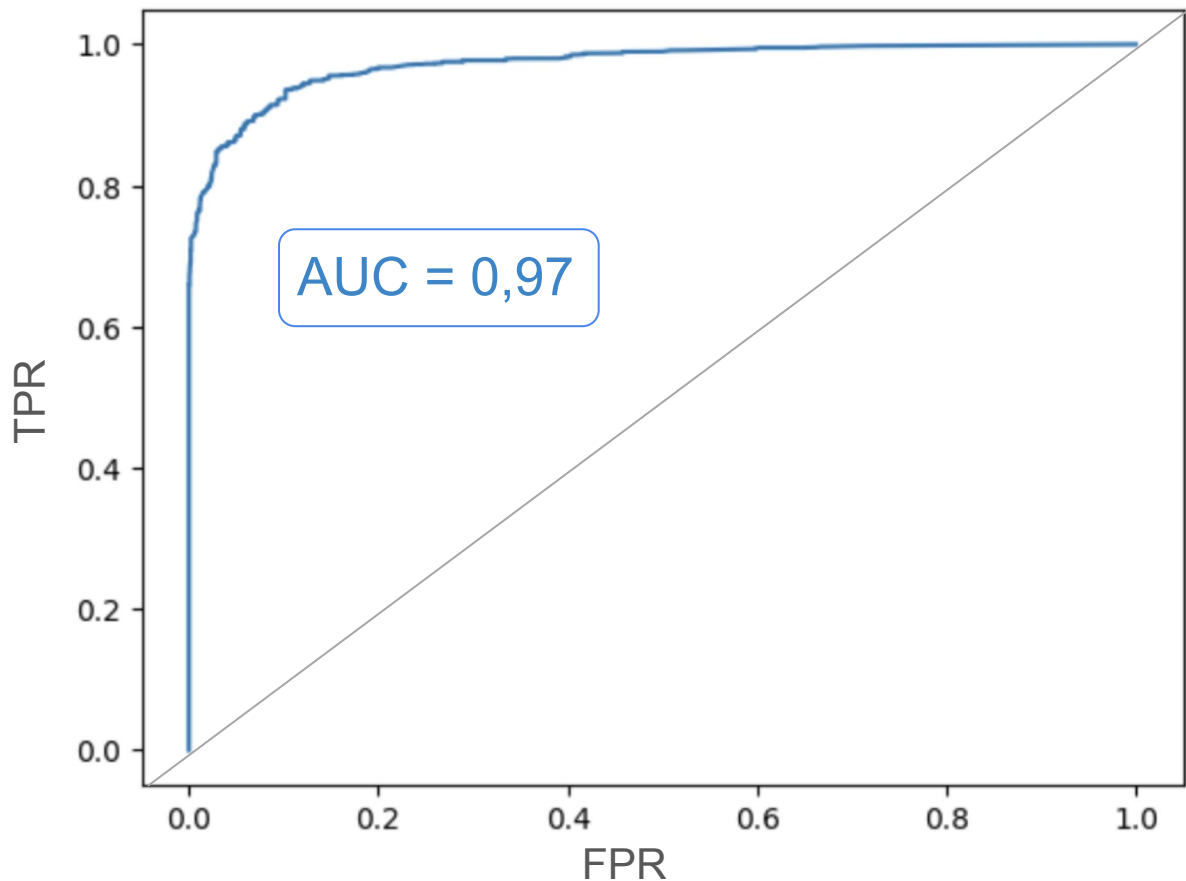
Градиентный бустинг LightGBM (Python)

- высокая точность
- больше предикторов
- сложность интерпретации



Предварительная оценка
экономического эффекта

Достигнутые метрики



Accuracy = 0.92

Precision = 0.90

Recall = 0.89

FNR = 0.11

FPR = 0.06

F1 = 0,90

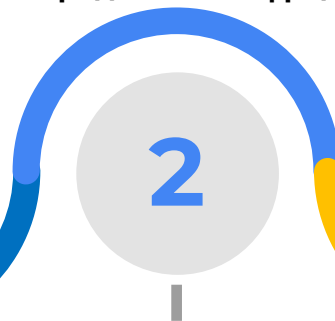
Оценка экономического эффекта

**20% - локальный уровень
фрода**



Полученные в результате данного исследования значения потенциального фрода составляют около 33%. Таким образом при значении $\text{Pres} = 0,89$, внедрение модели позволит дополнительно сэкономить сумму в размере 9,3% от зафиксированных страховых случаев.

**6 290 эскудо – средняя сумма
выплат
10 000 издержки за
ложноположительно
определенный фрод**



Введение машинной проверки страховых случаев приведет к ложноположительному обнаружению фрода в размере 6,7% от легальных транзакций, которые составляют как минимум 67% от всего количества случаев, т.е. 4,5% от всех страховых случаев.

**558 000 – общее количество
страховых случаев в год**



Таким образом, даже в случае максимальной доли проигранных исков при ложноположительном определении фрода, экономический эффект составит:

$$0,093 \times 558000 \times 6290 - 0,045 \times 558000 \times 10000 = 75,3 \text{ млн эскудо}$$

Искали мошенничества:



Малинова Ксения -
“белая экспатка”



Трогова Алина -
“капитан очевидность”



Горелик Борис -
“директор по
безопасности”



Мир Илья -
“доктор кабэ-верде-
форте”