

Предотвращение мошеннических убытков в личных видах: технологии черных списков

Мосесов Михаил Константинович
Руководитель управления развития андеррайтинга
ООО «Абсолют Страхование»

08.04.2025

Как отсекать мошенников в личных видах, если они обходят черные списки, а страхование в видах НС и ВЗР все привлекательнее и привлекательнее для множественного урегулирования?

Самый первый и главный инструмент отсеечения несегмента для страховщика на этапе предстрахового скоринга – это механизм черных списков.

Проверка происходит по сущностям/комбинации сущностей:

ФИО

Дата рождения

Документы

Адрес

Прочие параметры/индикаторы

Самый первый и главный инструмент отсека не сегмента для страховщика на этапе предстрахового скоринга – это механизм черных списков.

Проверка происходит по сущностям/комбинации сущностей:

ФИО

Дата рождения

Документы

Адрес

Прочие параметры/индикаторы

На языке программирования — строки:

Оформление полиса страхования от несчастного случая

1. Расчет > 2. Ввод данных > 3. Оплата > 4. Получение по e-mail

Фамилия страхователя

Чмелев

Выберите вариант или продолжите ввод

- Чмелева
- Чмелев**
- Чмелевская
- Чмелевский

Поле «Фамилия страхователя» является строкой, а набор букв ч-м-е-л-е-в представляет собой набор символов, в которых мы попробуем совершить подлог за счет изменения значений.

Расстояние Левенштейна

в компьютерной лингвистике — это минимальное количество операций подбора, удаления или замены нескольких символов, образующих полноценную строку.

Представим уравнение:

$$d(u, v)$$

где d – это мера различия символов (подбора/удаления/замены), требуемая для превращения u в v :

Расстояние Левенштейна

в компьютерной лингвистике — это минимальное количество операций подбора, удаления или замены нескольких символов, образующих полноценную строку.

Представим уравнение:

$$d(u, v)$$

где d – это мера различия символов (подбора/удаления/замены), требуемая для превращения u в v :

Допустим,

что расстояние $d = 1$

В нашем случае $v = \text{Чмелев}$

В свою очередь, u может послужить набор/кортеж = Чмеев, Челев, Чиелев, Чмалев, Чмулев, Чмелеев...

Для примера, на расстоянии $d = 2$,

$u = \text{Члев, Чмев, Чулев, Чиилев...}$

Данный пример помогает избежать аргумента «ошибся в букве», нужно применять в каждой отдельной строке для более точного семантического результата, можно остановиться на дистанции $d = 1$, в ФИО хватает 25-30 повторений.

Прототип реализован на MySQL.

Расстояние Левенштейна может быть реализовано как в режиме online, так и стать основой классификатора в виде машинного обучения: обучаться на строках ФИО+ДР, Адрес и Документы — в любом полезном для Ваших целей порядке.

Прототип реализован на MySQL.

Расстояние Левенштейна может быть реализовано как в режиме online, так и стать основой классификатора в виде машинного обучения: обучаться на строках ФИО+ДР, Адрес и Документы — в любом полезном для Ваших целей порядке.

Рекомендуем использовать дистанцию Левенштейна **d=1** для каждого параметра автономно, в том числе для документов, и отдельно производить настройку на соответствие города и адреса: исключить город и смотреть предметно на адрес.

Прототип реализован на MySQL.

Расстояние Левенштейна может быть реализовано как в режиме online, так и стать основой классификатора в виде машинного обучения: обучаться на строках ФИО+ДР, Адрес и Документы — в любом полезном для Ваших целей порядке.

Рекомендуем использовать дистанцию Левенштейна **d=1** для каждого параметра автономно, в том числе для документов, и отдельно производить настройку на соответствие города и адреса: исключить город и смотреть предметно на адрес.

Предпочтительно использовать алгоритм как базу данных по аналогу ЧС, содержащую вариации из ЧС, снижая риск отказа от сегмента. Уравнение прекрасно тем, что мы знаем исходное значение **v**.

Прототип реализован на MySQL.

Расстояние Левенштейна может быть реализовано как в режиме online, так и стать основой классификатора в виде машинного обучения: обучаться на строках ФИО+ДР, Адрес и Документы — в любом полезном для Ваших целей порядке.

Предпочтительно использовать алгоритм как базу данных по аналогу ЧС, содержащую вариации из ЧС, снижая риск отказа от сегмента. Уравнение прекрасно тем, что мы знаем исходное значение **v**.

Рекомендуем использовать дистанцию Левенштейна **d=1** для каждого параметра автономно, в том числе для документов, и отдельно производить настройку на соответствие города и адреса: исключить город и смотреть предметно на адрес.

Для пополнения знаний о значении **v** нам нужно больше данных об обогащении и мошенничестве в личных видах, нужна единая база данных с индикаторами по аналогии баз данных ОСАГО, КАСКО и жилья в НСИС и БСИ.



АБСОЛЮТ СТРАХОВАНИЕ.
ЗАЩИТА ВАШИХ УСПЕХОВ

+7 (495) 025-77-77

ул. Ленинская слобода, дом 26, стр. 4,
г. Москва, 115280

www.absolutins.ru



Мосесов Михаил Константинович
Руководитель управления развития андеррайтинга
ООО «Абсолют Страхование»