

Яндекс

Служба ассессоров

Служба ассессоров

- 9 стран
- 150 регионов
- Тысячи ассессоров

Про каждого ассессора мы знаем

- Типы заданий, которые он умеет делать
- Точность по каждому типу заданий
- Производительность и загрузку
- Пол, возраст, регион, периоды активности
- Время, за которое ассессор ставит оценку
- Правильные/неправильные оценки
- и т.д.

Рекомендатор ассессоров

Задачи сервиса (lvl 1)

- Отображать на схематичной карте мира различные параметры ассессоров
- Возможность фильтрации ассессоров по заданным критериям:
 - ассессоры онлайн
 - (не)делающие определенные типы заданий
 - из определенного региона
 - опытные/новички
 - с положительной/нулевой кармой
 - делающие в среднем не меньше M оценок в месяц
 - ...

Задачи сервиса (lvl 80)

- Найти зависимости между качеством выполнения определенных заданий и другими факторами (точностью в других заданиях, производительностью, социально-демографическими параметрами и т.д.) с помощью статистических методов или машинного обучения
- На основе выделенных зависимостей построить рекомендательный сервис «Для выполнения задания с заданными критериями вам подойдут следующие ассесоры»

В команду проекта нужны

- Два-три человека, желательно знакомых с ООП, maven, git, современными средствами разработки.
- Из них:
 - Один человек, знакомый с frontend-разработкой или желающий изучить соответствующие технологии (javascript, html5, css, ...)
 - Один-два человека для backend-разработки (java)

Зачем? Во-первых, это красиво



А ещё вы...

- Получите опыт работы с современными средствами визуализации данных (например, d3js)
- Построите веб-сервис «с нуля»
- Попробуете свои силы в разработке рекомендательной системы
- Освоите новые технологии и улучшите свои навыки работы с уже известными

Определение читеров

Зачем?

- Не все ассесоры работают добросовестно. Некоторые пытаются «играть не по правилам», ищут способы обмануть систему.
- Ассесоров много, в ручном режиме возможен мониторинг только самых простых нарушений (регулярное невыполнение нормы, точность ниже пороговой, и т.д.).
- Сложные случаи недобросовестности требуют наличия автоматических методов контроля работы ассесоров.

Задачи проекта

- Сформулировать и проверить гипотезы о том, как поведение читера может отличаться от поведения нормального работника, например: скорость оценки читера vs скорость оценки нормального работника, распределение оценок читера vs распределение оценок нормального работника и т.д.
- Написать алгоритм обнаружения читеров, используя выделенные статистические критерии, возможно с использованием машинного обучения

Команда проекта

- Один человек, хорошо знакомый с методами математической статистики, знакомый с методами машинного обучения и навыками программирования на java (последнее желательно, но не обязательно)

Зачем?

Получить опыт анализа данных математическими методами, попрактиковаться в использовании методов машинного обучения

Исследование «Сами мы не местные»

Стоит ли ассессорам оценивать запросы только из своего региона?

Главный вопрос исследования

Как распределять задания между ассессорами, чтобы получать наиболее точные оценки?

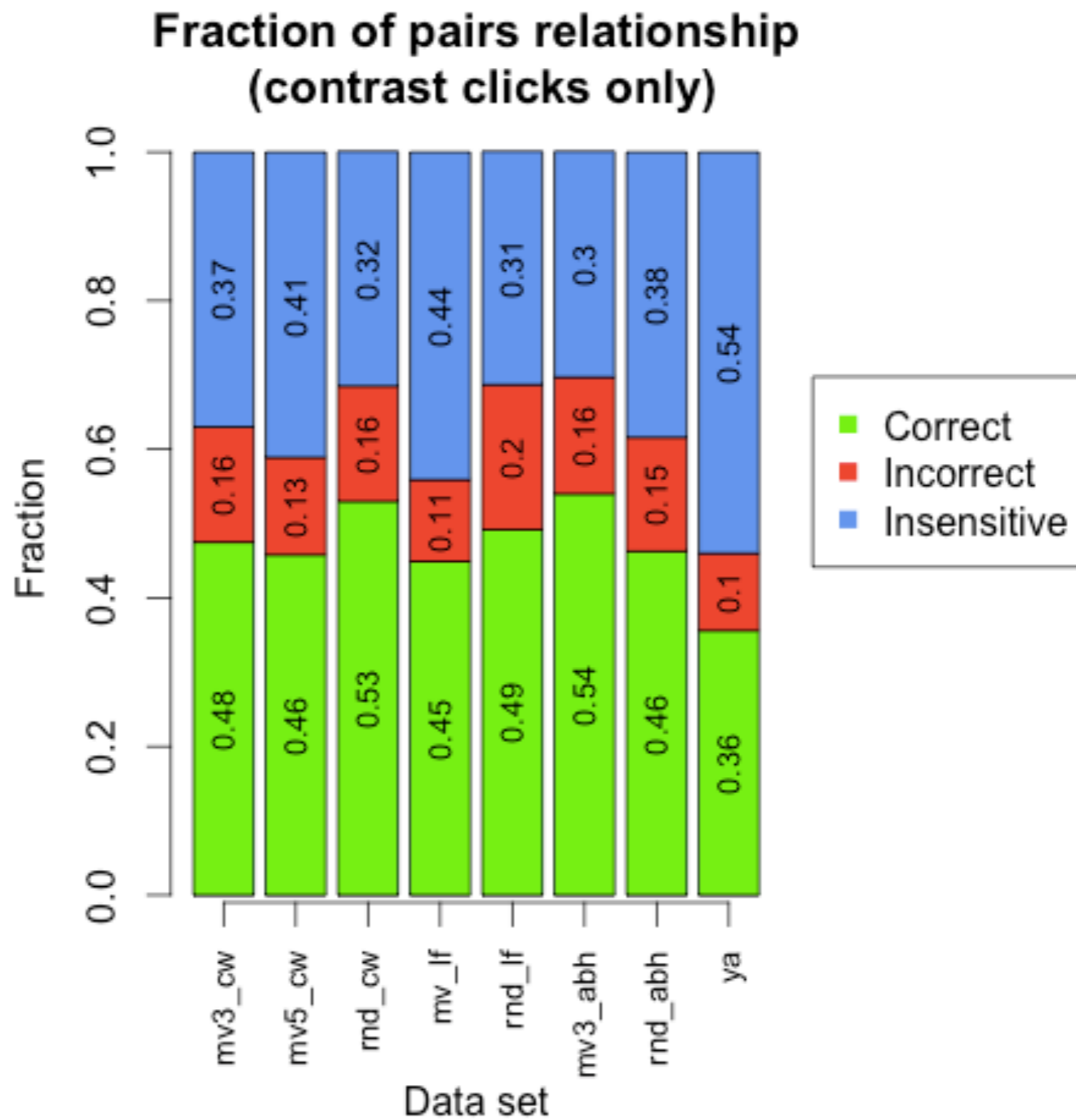
По мотивам исследования

An Analysis of Systematic Judging Errors in Informational Retrieval

Задача

Проверить гипотезу о том, что оценки ассессоров-аборигенов по запросам из данного региона лучше соотносятся с кликами, чем оценки произвольных ассессоров

Как часто оценки противоречат кликам?



Команда проекта

- Нужен один человек со знанием основ статистики и знанием любого языка программирования, достаточного для обработки текстовых файлов

Чему научитесь

- Организовывать эксперимент
- Работать с живыми данными
- Анализировать и интерпретировать результаты эксперимента
- Познакомитесь с передовыми научными исследованиями в области экспертной оценки

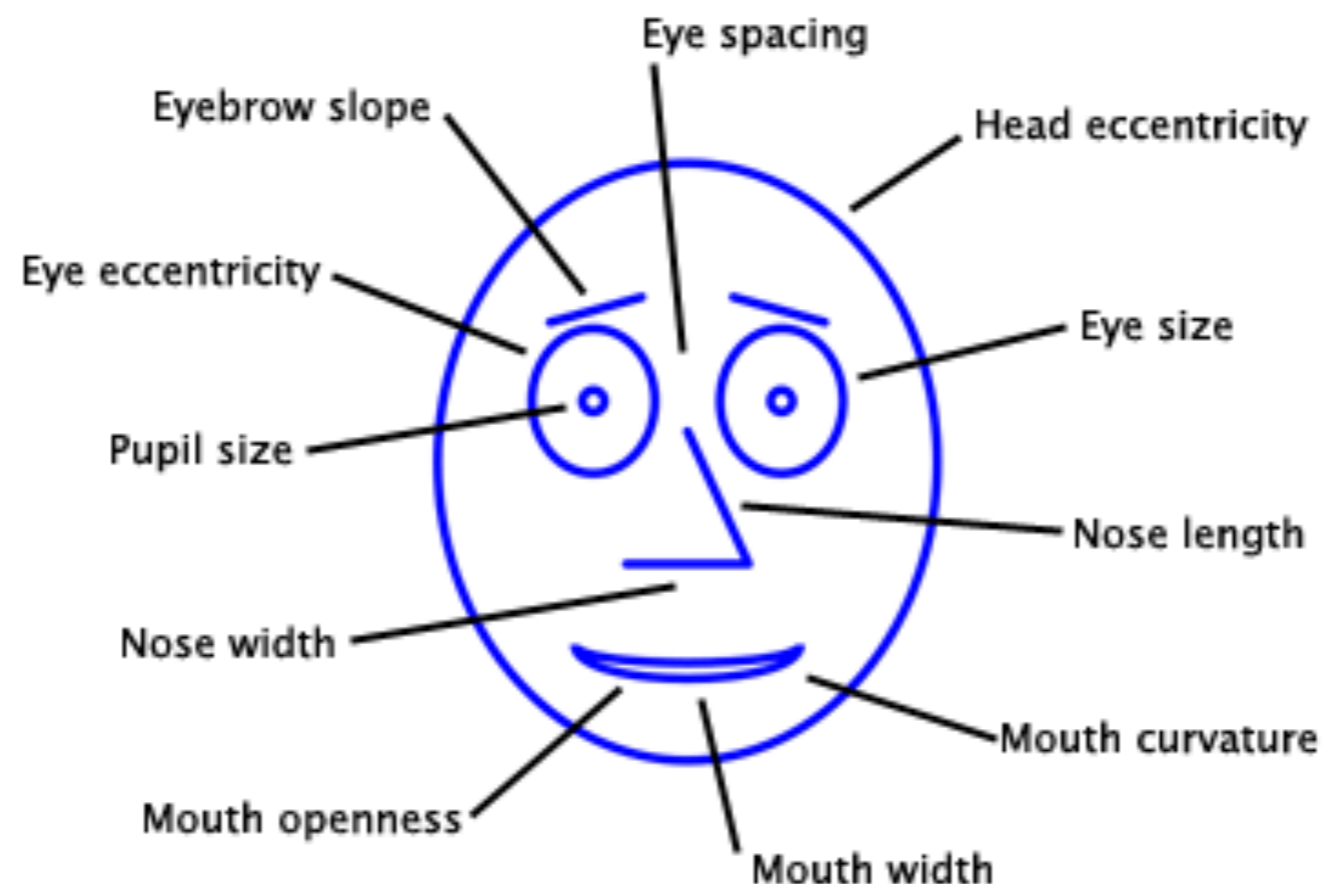
Лица Чернова

Задача

- Научиться сравнивать ассессоров между собой
- Ассессоров много, информации о них тоже много, традиционные способы представления информации неудобны для данной задачи

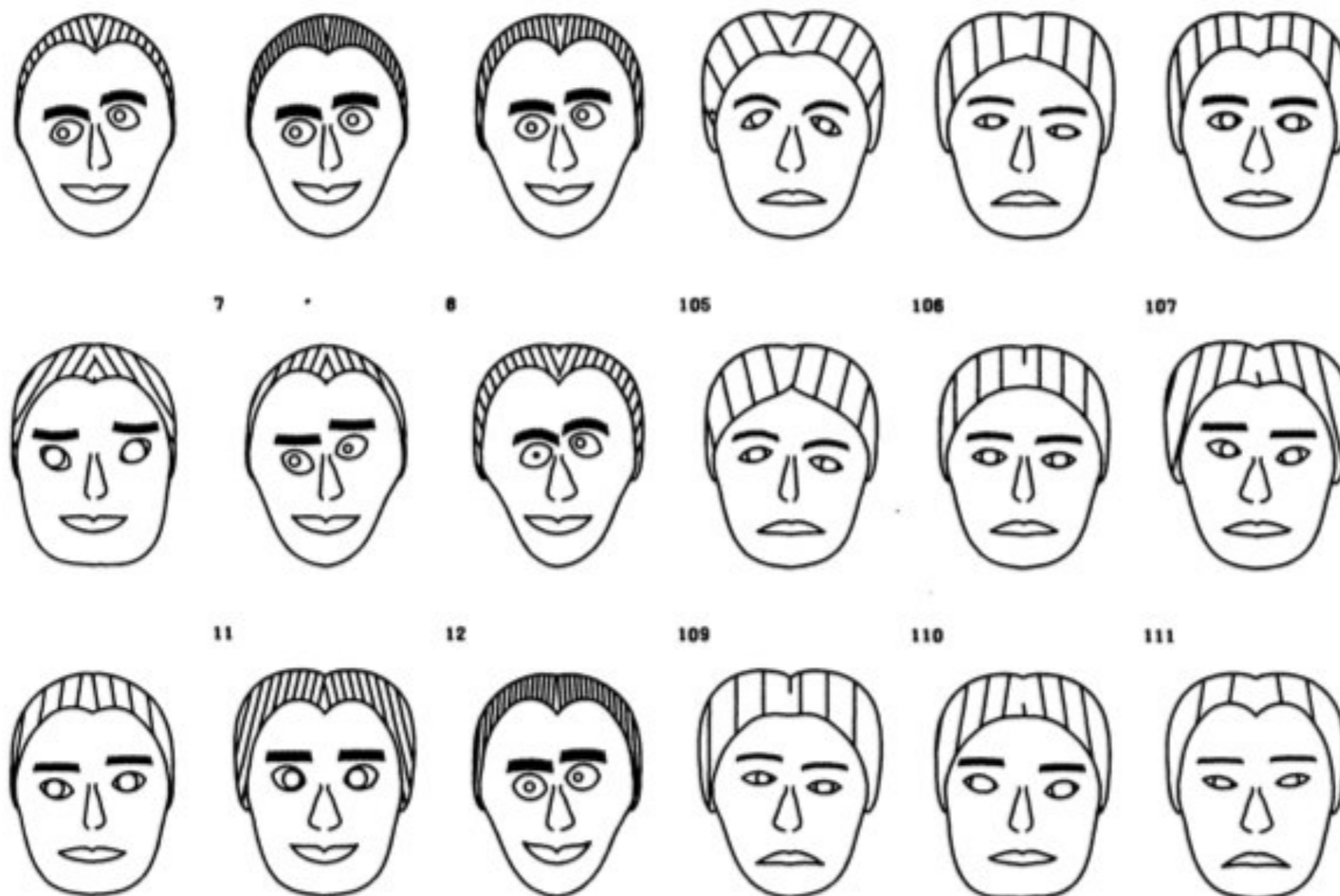
Решение

Германом Черновым был создан метод представления множества параметров в виде легко воспринимаемых образов – человеческих лиц. Эти образы позволяют видеть сразу множество параметров и замечать отклонения от нормы.



Классический пример

Анализ настоящих и поддельных банкнот по размерам границ, отступов и диагоналей



Требования к сервису

1. Конфигурация связей исследуемых параметров с чертами лица
2. Просмотр лиц в виде галереи
3. Выбор ассессоров для сравнения
4. Среднее лицо группы ассессоров
5. Удобство использования и приятный вид :)

Команда проекта

- В команду нужны два человека, желательно знакомых с ООП, maven, git, современными средствами разработки
- Один человек для разработки frontend-части, знакомый или желающий изучить соответствующие технологии (javascript, html5, css, svg, etc...)
- Один человек для разработки backend-части на java

Почему именно этот проект?

- Опробуете на практике необычный метод представления информации
- Разработаете веб-сервис «с нуля»
- Познакомитесь с новыми технологиями и улучшите навыки работы с уже известными

Записывайтесь в наши проекты!