

Модель профориентационного теста

Всероссийская студенческая конференция
«Студенческая научная весна»

Кирьянова Д.Р., ФН1-41Б
Хаддад Н.С., СГН3-64Б

Научные руководители: Гордеева Н.М., Труфанов Н.Н.

МГТУ им. Н.Э. Баумана

25 апреля 2024

Цель и задачи проекта

Цель проекта

Исследовать возможности использования “косинусной близости” для профориентационного тестирования.

Задачи проекта

- ❶ Составить “портрет” некоторых профессий (выбраны математики, физики, медики), для чего необходимо:
 - опросить представителей разных профессий через форму на сайте,
 - проанализировать ответы и определить веса вопросов,
 - создать шаблонные образцы, характеризующие выбранные профессии.
- ❷ Векторизовать ответы.
- ❸ Реализовать программу, вычисляющую “схожесть” ответов испытуемых с ответами, характерными для выбранных профессий.
- ❹ Проанализировать результаты.

Немного про сам тест

- Составляем 36 утверждений с ответами “да – нет”.
- Выбираем группы профессионалов (врачи, математики и физики).
- Представителей профессий просим пройти тестирование (каждой группе выделена своя роль).
- Желающие определить свой тип также отвечают на вопросы, по отдельной ссылке, но их ответы сравниваются с “шаблонными”.
- В качестве ответа выводятся значения косинусного сходства с каждой профессией.

Всего тест прошли: 14 медиков, 4 математика и 4 физика.

Платформа для сбора ответов

Questions

[Edit Questions](#)

Я стараюсь поддерживать порядок там, где это в моих силах.
Большинство успешных людей являются хаворорками.
В 21 веке странно употреблять в пищу братаев неманик, ни может позволить себе вегетарианство.
Я могу рассказать несколько стихотворений наизусть. Люблю поэзию!
Обычно я принимаю решения не колеблясь.
Лето лучше зимы.
Если в доме есть собака (питомца), то хозяева точно хорошие люди.
За последний год ни одной художественной книги, только по работе..
В политику идут те, кто не смог научиться ничему другому.
Каждое утро – радость для милуаха.
Я стараюсь равномерно распределить работу, не загибая до последнего.
Успешность любого человека на 99 процентов зависит от поддержки семьи и только на 18 процентов от него самого.
Все уже привели, что основная работа делается за день до указанного срока, надо это учитывать.
После 18 часов я стараюсь не есть (или хотя бы не объедаться).
Современная литература интереснее классической. (Классическая литература переоценена)
Сконцентрироваться для меня не проблема.
Дресс-код – это хорошо, главное, чтобы он не был слишком строгим.
Драться не очень хорошо, но каждый мужчина должен уметь драться.
Честно признаюсь, «Война и мир» мною не прочитана даже наполовину.
Мне интереснее посмотреть кино, чем театральную постановку.
Каждый человек имеет право на ошибки, но почему-то не каждый имеет право на критику.
Неспециальная лексика на работе для меня неприемлима.
Спортсмен (или физкультурой) я стараюсь заниматься не реже раза в неделю.
Телевизор смотрят люди от 40 лет и старше.
Чаще я смотрю контент в интернете, чем читаю книги.
Фразу «отрицательный результат – тоже результат» придумали неудачники.
Физкультуру в вузах надо отменить. Студенты взрослые люди, сами решат, что им делать.
Обожаю горные лыжи!
Машина в наше время – необходимость.
Вудуизм за ИИ.
Алкоголь – зло, но его отсутствие – беда.
Вудуизм за атомной энергетикой.
Без проблем сделаю прививку, если на то есть показания.
Кошек заводят старики или инфантильные люди.
Балет интереснее, чем опера.
Стараюсь избежать прививок, рассчитывая на собственный иммунитет.

[Download Answers](#)

Рис. 1. Админ-панель

Привет, спасибо, что согласились пройти наш опрос.

Пожалуйста, введите ваше имя:

Введите ваше имя

Выберите свою роль:

[Начать тестирование](#)

Вопрос 1 из 36

Я стараюсь поддерживать порядок там, где это в моих силах.

[Назад](#) [Вперед](#)

Рис. 3. Пример вопроса

Обработка полученных ответов

На этапе предварительной обработки данных удалим вопросы, на которые представители всех профессий ответили одинаково, так как они не являются информативными и не влияют на конечный результат.

Вычислим средние значения по профессиям. Для k -ой профессии формула для центроида класса имеет вид

$$\vec{\mu}_k = \frac{1}{n} \sum_{i=1}^n \vec{x}_i,$$

где n — количество представителей в k -ой профессии.

Введём весовые коэффициенты.

- для наиболее отличающихся ответов $\omega = 1$,
- для остальных $\omega = 0.1$.

Косинусная схожесть двух векторов \vec{a} и \vec{b} вычисляется по формуле

$$\cos(\widehat{\vec{a}, \vec{b}}) = \frac{(\vec{a}, \vec{b})}{\|\vec{a}\| \cdot \|\vec{b}\|} = k.$$

Получим коэффициенты k_i для каждой профессии и сравним их с коэффициентами K_i , полученными после введения весов

$$K = \frac{\omega \cdot (\vec{a}, \vec{b})}{\sqrt{\omega \cdot \vec{a} \cdot \vec{a}} \cdot \sqrt{\omega \cdot \vec{b} \cdot \vec{b}}}.$$

Результаты

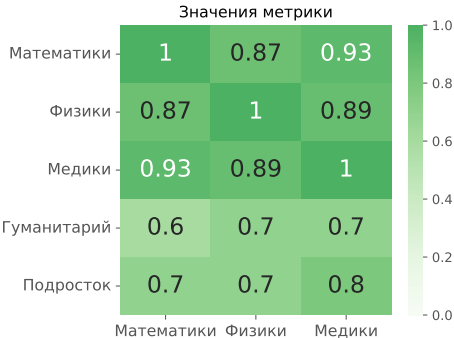


Рис. 4. Невзвешенные значения коэффициентов k

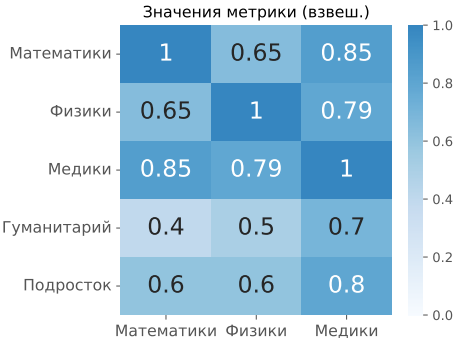


Рис. 5. Взвешенные значения коэффициентов K

- ❶ Рассмотренный метод работает, можно обращаться к психологам, чтобы улучшить качество вопросов.
- ❷ Увеличение размеров выборки данных делает модель точнее.
- ❸ Необходима проработка математических методов получения оценок весовых коэффициентов.