

Лебедев Юрий Сергеевич

магистрант

ОУП ВО «Академия труда

и социальных отношений»

г. Москва

МЕТОДЫ РАНЖИРОВАНИЯ ПРЕДЛОЖЕНИЙ НА ЭТАПЕ «ГОЛОСОВАНИЯ» КРАУД-ПРОЕКТА

***Аннотация:** в представленной статье анализируются методы ранжирования предложений технологии краудсорсинг на этапе «Голосование». Будет дана краткая характеристика тезиса ранжирование, также определены объекты ранжирования, приведены примеры объектов. Затем будут выделены основные задачи ранжирования в современном мире, также будет приведено краткое описание следующих методов ранжирования: метод прямого ранжирования, метод парных сравнений Терстоуна, экспертный метод, метод обратного рейтинга, метод арифметической прогрессии.*

***Ключевые слова:** ранжирование, краудсорсинг, методы, голосование, Лайкерт, шкала Гуттмана, Терстоуна.*

Ранжирование является процессом сортировки группы элементов по соображениям релевантности. Принято подразумевать под релевантностью отношение по рангу между объектами. На последнем этапе проекта краудсорсинга «Голосование» получается объемный массив из предложений, отобранных по ходу проекта. На этапе «Голосование» поставлена задача ранжировать массу идей и выделить лидирующие.

Чаще всего ранжирование – это процедура упорядочения любых объектов по возрастанию или убыванию одного из их свойства при условии, что это свойство имеется у объекта.

Под объектами ранжирования подразумевают объекты, перед которыми стоит задача упорядочивания. Эти объекты могут обладать совершенно разными свойствами.

К примеру, можно ранжировать граждан города Москвы по уровню заработной платы, по их отношению к чему-либо, по престижности района, где эти граждане проживают. Также можно ранжировать государства по уровню жизни, площади, ВВП и т. д. Так же, есть возможность ранжировать пользователей информационной системы по уровню их активности, рейтингу и другим показателям.

Основание ранжирования – это свойство, по которому объекты упорядочиваются. Результатом процесса упорядочения объектов является ранжированный ряд, в котором каждому объекту приписывается ранг, он же место в этом ряду. Число мест и, соответственно, число рангов равно числу объектов.

Задача ранжирования сегодня очень востребована: сортировка веб-страниц согласно заданному поисковому запросу, персонализация новостной ленты, рекомендации видео, товаров, музыки, подсчет голосов в опросе и т. д. Существуют специальное направление в машинном обучении, которое занимается изучением алгоритмов ранжирования способных самообучаться – обучение ранжированию (англ. learning to rank).

Ранжирование является приемом анализа, приемом решения исследовательских задач.

В краудсорсинговой информационной системе на этапе «Голосование» существуют две задачи, напрямую связанные с ранжированием предложений по степени их популярности среди пользователей системы:

1. Первая задача: как грамотно организовать пул предложений для процесса голосования, которое будет доступно пользователю.
2. Вторая задача: получение результатов ранжирования ряда предложений пользователей, по средствам подсчета голоса отданных на этапе «Голосование» за те или иные предложения.

Применив подходящий метод к первой задаче, возможно упростить процесс второй задачи до подсчета общей суммы голосов, отданных за предложения.

В информационной системе задействованы пользователи совершенно разных категорий: по специальности, возрасту, уровню жизни и т. д. При этом

придерживаемся идеи, что перед пользователем имеется большое количество предложений совершенно различной тематики и качества. Оценить все предложения, демонстрируемые ему, пользователь не может потому, что перечень предложений для оценки массивен.

Существует различие между ранжированием и измерением по порядковой шкале. Оценка получается либо на основе приписывания «цифры» вариантам ответа, либо по логическому квадрату, либо по шкале Лайкерта, либо по шкале Гуттмана.

Выбор алгоритма ранжирования процесс сложный. Необходимо иметь четкое понятие требуемого результата, оценивать качество работы алгоритмов.

Рассмотрим пример, который может возникнуть в системе: пользователь, открыв первые два предложения, может быть полностью удовлетворен содержанием полученных в этих двух предложениях. Пользователь произведет положительную оценку первых двух предложений, и, вероятность, что он откроет предложение под номером 3 и выше, крайне мала.

Подобные модели поведения пользователя, где изучение предложенных ему элементов происходит последовательно, и вероятность просмотра элемента зависит от релевантности предыдущих, называются каскадными. Сегодня существует несколько известных методов ранжирования, именно о них далее пойдет речь. Ранжирование может осуществляться несколькими способами, но в данной краудсорсинговой системе ранжирование происходит на основе данных арифметической прогрессии.

На данный момент в системе пользователю демонстрируется поочередно, группа из 20 предложений за группой. Получается, что просмотр каждого элемента независим от просмотров других элементов – своего рода «наивность». На практике же, элементы зачастую просматриваются пользователем поочередно, и то, просмотрит ли пользователь следующий элемент, зависит от его удовлетворенности предыдущими.

Пользователь на этапе «Голосование» имеет возможность только поставить одно значение предложению – положительное.

Основанием ранжирования в этом случае является сумма голосов пользователей. При этом никак не учитывается опыт пользователя на прошедших краудсорсинговых проектах, его активность и прочие значения.

После проставления оценки пользователем, в краудсорсинговой системе происходит простое арифметическое суммирование положительных голосов индивидуально у каждого предложения, а затем вычисление места в ранжированном ряду.

По нашему мнению, этот вид ранжирования не только недопустим, но и практически приводит к большим погрешностям на этапе.

Существует на данный момент множество методов ранжирования данных. Каждый из методов ранжирования обладает своей собственной спецификой, своими недостатками и преимуществами. Одни из самых востребованных методик ранжирования данных, в том числе и используемый метод ранжирования:

1. Метод прямого ранжирования.
2. Метод парных сравнений Терстоуна.
3. Экспертный метод.
4. Метод обратного рейтинга.
5. Метод арифметической прогрессии.

Далее проанализируем их преимущества и недостатки.

Метод прямого ранжирования является наиболее простым, с технической точки зрения, методом оценки. Суть данного метода заключается в присвоении ранга в соответствии со степенью популярности у пользователей системы. В системе на этапе «Голосование» возникает задача ранжирования предложений.

Пользователи проставляют положительные оценки выбранным предложениям, со значением на их выбор.

Сумма всех голосов предложений является первичными данными для вышеописанного метода. Условием в данном случае является построение ряда по значению голосов от большего к меньшему.

Метод парных сравнений Терстоуна этот метод был впервые введен в обращение Луи Терстоуном и был использован для ранжирования преступлений по

степени серьезности и для ранжирования различных национальностей по предпочтительности с точки зрения дружеских отношений. Метод парных сравнений основан на попарном сравнении объектов ранжирования по заданному основанию.

Процедура сбора данных происходит следующим образом. На отдельные карточки заносятся названия объектов ранжирования. Карточки перетасовываются, и пользователю предъявляется первая пара с вопросом: *«Оцените пожалуйста, какое предложение актуальнее и больше направлено на решение поставленной задачи перед проектом?»*. Затем предъявляется вторая, третья пара и т. д.

Имеет место быть такой метод как *Экспертный метод*. Суть экспертного (лат. expert – «опытный») метода заключается в использовании мнения группы квалифицированных экспертов. Предполагается, что качество мнения группы экспертов выше, чем качество отдельного эксперта или же группы менее квалифицированных людей. Подобный метод часто используется для оценивания сложных задач с высоким требованием качества.

Метод обратного рейтинга – идея «Антирейтинга». Суть ее заключается в том, что пользователю предлагают выбрать не лучший вариант предложения, а отсеять худший вариант предложения. Данная идея может быть наложена как на метод парных сравнений Терстоуна: *«Оцените, пожалуйста, какое предложение менее подходящее для решения поставленной задачи перед проектом?»*, так и на метод прямого ранжирования предложений, где от пользователя потребуется ставить минимальные оценки неподходящим предложениям, по мнению самого пользователя.

Слабой стороной данного метода является его нераспространенность, и вследствие чего, его нестандартность для пользователей. Потому данный метод рассматривать категорически нельзя.

Метод арифметической прогрессии. Перед пользователем стоит задача оценить предложение или пропустить его. В случае положительной оценки пользователь проставляет оценку «+1» и у предложения становится количество голосов N (количество голосов до демонстрирования) + значение 1. В случае если

пользователь не оценивает предложение, демонстрируемое ему, у предложения останется количество голосов равное количеству голосов до демонстрирования.

Список литературы

1. Ивушкина Н.В. Социальный эффект инвестиционных процессов. – 2014. – С. 5–7.
2. Интеграция науки, образования и производства – стратегия развития инновационной экономики: Материалы I Международной научно-практической конференции. – 2014. – С. 153.
3. Калянов Г.Н. Информационные технологии и управление предприятием / Г.Н. Калянов, В.В. Баронов, Ю.И. Попов, И.Н. Титовский. – 2016. – С. 328.
4. Кузьмин Е.И. Оценка социальной эффективности деятельности по продвижению: Справочник руководителя. – 2017. – С. 26–28.
5. Характеристика наиболее используемых методов анализа рисков 2005[Электронный ресурс]. – Режим доступа: <http://naukovedenie.ru/PDF/22ergsu313.pdf>
6. Ханова А. Социальный эффект: практика измерения [Электронный ресурс]. – Режим доступа: <http://www.cloudwatcher.ru/analytics/1/view/36/>