OpenSAP, Stepic:

Поощрения: Дополнительные баллы оценщику

**Coursera (самооценка, рубрика)**

Coursera **интегрировала систему коллегиальной оценки в учебную платформу для оценки и обеспечения обратной связи,** по крайней мере от 3 до 4 заданий. Coursera предоставляет учащимся дополнительного оценочную матрицу для улучшения результатов коллегиальной оценки. **Кроме того, учащиеся имеют возможность самостоятельно оценить себя** (Piech et al., 2013; Luo et al., 2014). Система коллегиальной оценки в Coursera включает в себя три основных этапа: 1) фаза представление, 2) фаза оценки, и 3) опубликование результатов (Coursera, 2015). До недавнего времени не было никаких достоверных доказательств о том, как коллегиальная оценка влияет на процесс обучения в Coursera.

В нескольких MOOCs предлагаемых Pennsylvania State University и размещенные онлайн на Coursera, учащиеся сообщили, что они не доверяют результатам коллегиальной оценки. Кроме того, они изложили некоторые претензии к коллегиальной оценке, такие как отсутствие коллегиальной обратной связи, точность и доверие (Suen, 2014).

**В дополнение к вопросам множественного выбора и компьютерного оценивания, коллегиальная оценка была использована в ряде курсов Coursera. Студенты сначала овладевают с оценочной рубрикой. Чтобы получить обратную связь о своей работе, студент должен оценить пять эссе.** Затем студент получает коллегиальные оценки от пяти других студентов. Используя алгоритмы машинного обучения**, Piech и соавт. оценили и с поправкой на грейдер предубеждений и надежностей [9].** Они демонстрируют значительное улучшение в точности оценивания на реальных данных свыше 63,000 коллегиальных оценок. Их модели вероятностные, так дают распределение убеждений над оценками (в отличие от простого балла) для каждого студента.

EDX (рубрики)

Коллегиальная оценка в EDX работает похожим образом как и в Coursera. Здесь учащиеся обязаны рассматривать несколько образцов (примеров) заданий, которые уже были оценены профессором, прежде чем оценить своих коллег. После того как учащиеся доказали, что они могут назначить оценки, аналогичные тем, которые даны профессором, им разрешается оценить работу друг друга и обеспечивать обратную связь, используя ту же самую рубрику (EDX, 2015).

PEAS новое расширение коллегиальной оценки для **EdX платформы Mooc с открытым исходным кодом [10].** Студенты заинтересованы в оценивании точно методом калибровки, который строит счет стимулирования, основанную на точность их оценивания. Для улучшения качества обзора, студенты делятся на группы в зависимости от этой оценки мотивации и каждое задание коллегиально оценено по одному студенту из каждой группы. Простая нормализация оценок также осуществляется для уменьшения погрешности в коллегиальном оценивании. Оценивания экспертом могут быть использованы для устранения расхождений в коллегиальном оценивании, а также обеспечить обучающие данные для алгоритмов машинного обучения для автоматического оценивания. Важным отличием нашей работы является то, что калибровка в PEAS только один раз, в то время PeerRank потенциально использует несколько раундов согласования оценок.

Мы можем сделать определенные выводы о наиболее эффективных методов коллегиальной оценки в bMOOCs следующим образом:  
**Время**: Оптимальное обратная связь должна быть обеспечена в начале процесса оценки, с тем чтобы дать учащимся возможность среагировать и улучшить свою работу.  
**Анонимность**: важным аспектом коллегиальной оценки является обеспечение анонимности обратной связи. Таким образом, рецензенты могут обеспечить критическую обратную связь и сортировки без учета факторов, например межличностные смещения дружба или личные антипатии.  
**Доставка**: Косвенное обратной связи обеспечивает более эффективные результаты оценки в качестве учащиеся чувствуют себя более комфортно, чтобы дать честную обратную связь без влияния со стороны сверстников.  
**Peer оценивание**: коллегиальная оценка должна быть только частью итоговой оценки в целях обеспечения достоверности результатов оценки.  
**Несколько рецензентов**: Результаты оценки могут быть более точными и достоверными, когда учащиеся получают обратную связь от нескольких рецензентов, а не из одного. Таким образом, учащиеся имеют возможность получить разностороннее обратную связь в их работе.  
**Цикл оценивания**: Наличие нескольких итераций обратной достижения лучшего результата обучения, а учащиеся могут задуматься о работе назначение несколько раз.  
**Роль учителя**: учителя должны все еще принимать активное участие в процессе коллегиальной оценки, определяя оценки рубрики, обеспечивая примеры решения и проверки результатов коллегиального обзора. Они также могут помочь в развитии навыков рецензирования.

На основании обзора литературы коллегиальной оценки и результатов обследования, мы получили набор требований для поддержки коллегиальной оценки в L2p-bMOOC, кратко излагаются ниже:

**Пользовательский интерфейс:** Интерфейс должен быть простым, понятным, и простой в использовании, но требует минимального ввода данных пользователем. Разработка интерфейса модуля должны принимать принципы юзабилити во внимание и пройти через участии процесса проектирования (Nielsen, 1994).

**Рубрики:** предоставить учащимся с гибкими рубриками задач-специфическими, которые включают описания каждого элемента оценки для достижения справедливой и последовательной обратной связи для всех участников курса.

**Управление:** коллегиальная оценка должно быть легко управлять. Модуль должен быть интегрирован в платформу с функциями для активации и деактивации.

**Масштабируемость:** Фундаментальное различие между MOOCs и традиционным классом есть масштаб учащихся. Следовательно, масштабируемость следует рассматривать в реализациях модуль коллегиальной оценки в L2p-bMOOC.

**Совместная оценка:** Обеспечить механизмы для процесса совместной оценки, который включает оценивание более чем одним отдельного участником.

**Двойной слепой процесс:** модуль коллегиальной оценки должен поддерживать процесс двойного слепого рецензирования (обзора). Ни авторы задания не знают личности рецензентов, и наоборот.

**Сроки:** Модуль коллегиальной оценки должен предоставить два срока для каждой задачи: срок подачи для учащихся, чтобы представить свои работы, а другой для фазы оценивания.

**PeerRank Метод коллегиальной оценки**

Мы предлагаем метод PeerRank для коллегиальной оценки. Это создает оценку для агента на основе оценок, предложенных агентов, оценивающих агента. Поскольку оценка агента является мерой их способности оценивать правильно, PeerRank метод весовых оценок по оценкам оценивания агента. Метод PeerRank также обеспечивает стимул**ы** для агентов оценивать корректно. Как оценки агента зависят от оценок оценивания агентов, так и эти оценки сами зависят от оценок других агентов, мы определяем метод PeerRank уравнением неподвижной точке, аналогичным PageRank способу или рейтингу веб-страниц. Мы выявляем некоторые формальные свойства метода PeerRank (например, он удовлетворяет аксиомам единодушия, без пустышки, без дискриминации и симметрично), обсуждаем некоторые примеры, сравниваем с соответствующей работой и оцениваем производительность на некоторых синтетических данных. Наши результаты оказались весьма перспективными, уменьшая ошибку в прогнозах оценки с коэффициентом 2 или более, во многих случаях более он естественным базовой среднего коллегиальных оценок

Оценка каждого агента строится из оценок агентов, **оценивающих** агента. Поскольку оценка является мерой их способности оценивать правильно, мы взвешиваем оценку агента данную другим агентом по их собственной оценке. Оценка агента, таким образом, есть средневзвешенная из оценок агентов, оценивающих агента.

**Количественные и порядковые**

**В количественной коллегиальной оценке или абсолютного суждении (Барнетт 2003 года; Паре и Joordens 2008), в коллегиальных оценках для заданий в виде явных численных оценок. Это количественное значение в настоящее время наиболее распространенным выбором для MOOCs.** Его цель, как правило, чтобы свести к минимуму отклонения абсолютной предсказания от какого-наземного контроля. Тем не менее, (Шах и др 2013 года;. 2014) утверждал, что иногда легче для не-экспертов в коллегиальном оценивании, чтобы сделать сравнение, чем дать абсолютные значения и, следовательно, порядковый оценку иногда может быть более эффективным, чем количественной оценки.

Цель порядковой коллегиальной оценки состоит в предсказании относительного ранжирования правильно

Количественные+порядковые:

Сначала генерировать попарно преференций от количественных оценок каждого сверстников грейдер, чтобы сформировать обучающие данные. Галстуки, исключены, поскольку они указывают на отсутствие предпочтения. Наша основная предположение, что абсолютные оценки грейдер на множестве заданий также отражают ее порядковые предпочтения.

количественные модели превосходят чистые порядковые, потому что они используют мелкозернистую числовую информацию, а не двоичные сравнения (лучше или хуже) в порядковом обстановке

В настоящее время, порядковые сравнения генерируется из количественных оценок. Если порядковые данные генерируются непосредственно, сочетая количественные и порядковые модели могут быть использованы для дальнейшего мелодии и усиливают друг друга, так что дальнейшее улучшение возможно.

**Вопросы калибровки**

Каждый студент имеет присущие уклоны, которые могут меняться со временем или зависеть от качества другого ответы оценивали. Студент может быть консервативным и всегда назначать умеренные оценки, в то время как другой студент может иметь тенденцию завышать оценки в крайности. Эти несоответствия трудно чтобы узнать, с размерами выборки размером до 5 или 10, которые можно было бы получить, как правило, в установке по экспертной оценке.

Кроме того, эти несогласованности изменяются со временем, что делает их труднее изучать и модель. На с другой стороны, такие несоответствия устраняются автоматически (по дизайну) в порядковом установки.

**Конкретная Конструктивная обратная связь**

Учитывая, что оценщики не являются экспертами, мы выдвигаем, что порядковое обратной связи, как показано на рис. 1, может часто может быть более желательным для студента по сравнению с комментариями по абсолютной шкале. Так как люди лучше в сравнении предметов, оценщик может быть в состоянии предложить более проницательные замечания по положительные и отрицательные аспекты решения студента, и сделать эти пункты более конкретным путем сравнения с другой solution.1 Такая точечная оценка будет также направлять студента в понимании как именно он / она может улучшить его / ее решение.

**Удобство оценки**

Проблема независимой оценки в MOOCs требует решения не только из-за несоответствий отсутствие опыта, но и с дефицитом ресурсов. Это практически неосуществимым, чтобы спросить каждого студента уделять слишком много его / ее времени и усилий в решении задач по экспертной оценке. Таким образом, интерес к максимизировать количество информации, которая может быть собрана в соответствии с ограниченным усилием студентами, и разработать процесс оценки сверстников, не усталость студентов-оценщики. С этой целью его достаточно хорошо известно [6, 7], что люди часто находят значительно проще, чем сравнивать счет. *В Настройку независимой оценки, оценщик может найти задачу обеспечения простого сравнения быть гораздо проще, чем предоставление точной балльной оценки.* Порядковое подход позволит затем для сбор большего количества оценок для того же уровня усилий, по сравнению с оценками количественных.