



<http://aptitude.w3c.fmi.uni-sofia.bg/>

НАУЧЕН ОТЧЕТ

По проект APTITUDE

Иновативна софтуерна платформа за анализи на големи масиви от учебни и игрови данни за ориентирана към потребителя адаптация на технологично подпомогнато обучение

Фонд „Научни изследвания“, МОН,

Конкурс за финансиране на фундаментални научни изследвания по обществени предизвикателства – 2018 г.

Номер на договор: ОПР03/29 от 13.12.2018г.

D1.1. Сравнителни анализи на SOTA за препоръки и адаптация на учебно съдържание и работен поток от учебни дейности.

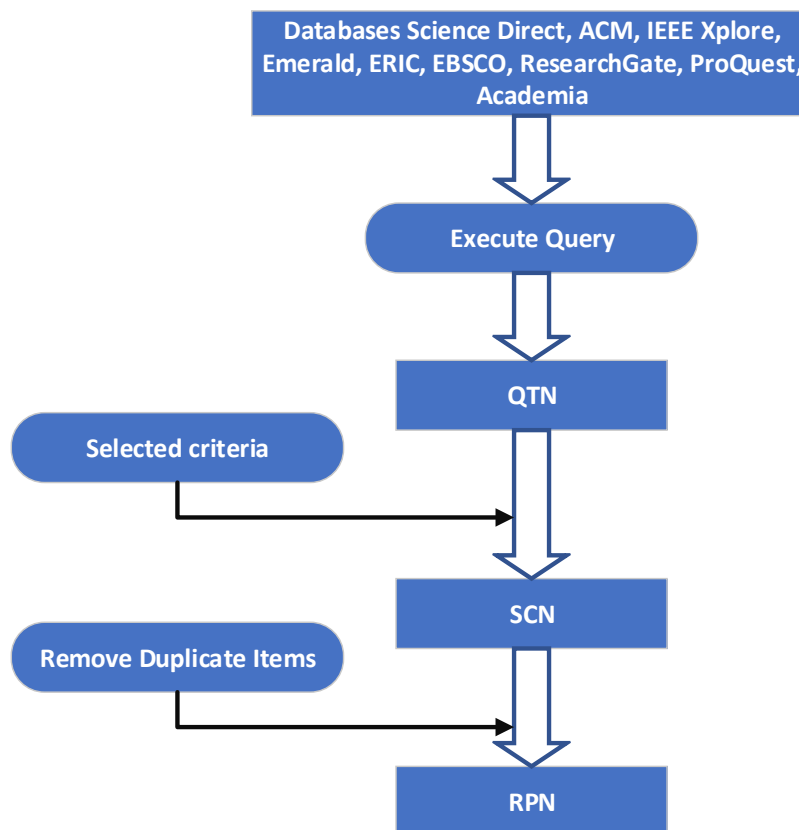


Съдържание

I. Методология на провеждане на проучване на състоянието на проблема (SOTA: state-of-the-art)	3
II. Съвременни тенденции при обработка и анализ на големи масиви от данни при системите за управление на обучението и образователните игри	4
1. Обработка и анализ на големи масиви от данни при системите за управление на обучението и образователните игри	4
2. Анализ на големи масиви от данни при обучението и игрите	9
III. Методите и техниките за препоръки и адаптация на учебно съдържание и работен поток от учебни дейности	12
1. Методите и техниките за адаптация на учебно съдържание и работен поток от учебни дейности	12
2. Техники за препоръка на учебно съдържание. Категории системи за препоръки.	16
Използвана литература	21

I. Методология на провеждане на проучване на състоянието на проблема (SOTA: state-of-the-art)

1. Проучване на съвременните тенденции се извършва по методологията показана на фиг. 1.



Фиг. 1 Методология на провеждане на проучване на състоянието на проблема

Изследването да използва дигитални библиотеки на Science Direct, ACM, IEEE Xplore, Emerald, ERIC, EBSCO, ResearchGate, ProQuest Index to Theses и Academia. В резултат се изпълнява съответната заявка по зададени ключови думи, като търсенето се осъществява по тема, анотация и ключови думи на статията. За съответните проучвания се изпълниха следните заявки:

- Проучване на обработка на големи масиви от данни при системите за управление на обучението и образователните игри: TITLE-ABS-KEY (data AND mining AND e-learning).
- Проучване на методите и техниките за препоръки и адаптация на учебно съдържание и работен поток от учебни дейности: TITLE-ABS-KEY (adaptation AND recommendation AND e-learning).

- Проучване на използването на анализа на големи масиви от данни при обучението и игрите: TITLE-ABS-KEY (learning AND analytics AND e-learning).

В резултат от изпълнението на заявката се получават в различните бази данни се получават определен брой статии - QTN (Query Total Number). След което се прилагат допълнителни критерии, за да се отстранят статиите, които не са свързани със съответния проблем и се получава нов брой на избраните статии - SCN (Selected Criteria Number). Например при търсенето с ключовата дума adaptation, която се среща и в machine learning, е необходимо да се убедим, че са отстранени всичките статии, които не са свързани с образователната област.

Третата фаза е да се прмахнат тези статии, които вече се дублират от другите източници и се получават брой статии за преглед и анализиране - RPN (Research Paper Number).

II. Съвременни тенденции при обработка и анализ на големи масиви от данни при системите за управление на обучението и образователните игри

1. Обработка и анализ на големи масиви от данни при системите за управление на обучението и образователните игри

Една от най-предизвикателните задачи за днешните организации са извличането на информация и стойност от данните, съхранявани в различни информационни системи във всяка една област. Тези данни, които са много сложни и твърде обемни, за да бъдат обработвани от традиционната системи за управление на база данни, се наричат големи масиви от данни. Големите данни е термин за група от масиви от данни, толкова масивни и усъвършенствани, че да станат трудно да се обработва с помощта на инструменти за управление на бази данни или съвременни обработка на приложения. Независимо дали става дума за ежедневни данни, бизнес данни или основни данни, ако те представляват масивна обем от данни, структурирани или неструктурирани, данните са от значение за организацията. Не само измеренията на данните са важни - това е как се използва от организацията за извличане на по-задълбочени прозрения, които могат да ги тласнат към по-добър бизнес и стратегически решения. Тези данни могат да бъдат използвани за определяне на качеството на изследванията, подобряване на процеса на процеса в организация и др.

Обработка и анализ на големи масиви от данни при системите за управление на обучението и образователните игри се развиват в четири основни модула с цел осигуряване на качествено образование и възможност за разработване на гъвкави учебни дейности:

1. Създаване и доставяне на електронно учебно съдържание
2. Оценяване
3. Комуникационни канали
4. Административно осигуряване

1.1. Създаване и доставяне на електронно учебно съдържание

Създаването и доставянето на електронно учебно съдържание е свързано с разработка на електронно учебно съдържание, които да поддържат средства за създаване на материалите за курс. От друга страна са необходими и среди, които да могат да се използват като хранилища (repository) за страници, мултимедия, видео клипове и др.

Ядрото на всяко едно електронно учебно съдържание е мултимедията, която има за основна цел да подобри интерактивността ученик-възпитател чрез систематизиране на отношенията между основните участници в обучението. Следователно, мултимедийната оценка на средата за електронно обучение се стреми да подпомогне организацията на дидактическия материал и да подобри системата за препоръки, за да направи учебния процес по-ефективен [21, 43, 46, 2].

В [83] е разработена система за оценка на обучението на учениците чрез анализ на факторите, който влияе върху тяхното поведение по време на използване на медията. За да се оцени нивото на обучение на ученика, е използвана техниката на дървото за вземане на решения. Целта е да подобри метода за оценка, за да подобри поведението на учениците във връзка с използването на учебната среда [83].

Търсенето на поведенчески модели на учениците при достъп и разглеждане на образователни ресурси е целта на статията [48]. Авторите извличат поведенчески модели, свързани с взаимодействието на ученика с образователното медии. За това авторите използват правилата за асоцииране и откриването на последователността в записите за достъп на ученика до учебния материал. Резултатът подчертава възприятието на ученика и идентифицира тенденциите по отношение на използването на образователни медии за учене [25].

Средите за електронно обучение генерират данни като регистрационни файлове за достъп, използваемост на медии, бележки за дейности и др. взаимодействия. Pascual-Cid et al. [59] разработи система за изследване и визуализиране на генерираните данни в тези среди, и се стреми да анализира тези данни, използвайки статистически техники за извличане на модели на поведение на ученика.

Данните се представят по интерактивен начин, като по този начин предоставят важни данни за оценките на преподавателя, за да се установи най-доброто преподавателски практики [13, 46, 71, 85] твърдят, че понастоящем сериозните игри са привлекли вниманието на преподавателите, представяйки допълнителна функция към учебния процес. Това изследване сравнява други методи за оценка като метод Delphi [3], АНР (Analytic Hierarchical Процес) [6] и метода на Fuzzy оценка [41]. Авторите предлагат метод на машинно обучение, базиран на изкуствените невронни мрежи като алтернатива на модела за оценка, използващ образователния игри. Моделът за оценка използва извадки за обучение, за да намали влиянието на субективните фактори и да предостави точни показатели за автоматично оценяване на образователните медии. В [66] се стремят да предоставят полезна информация за оценка на нивото на интерес на учениците към образователните медии, използвайки данните си за достъп до учебни материали (медии). Данните за достъп се групират и анализират, за да се разпознае поведението модели,

например, идентифицират дали инструкциите са били неадекватни или недостатъчни, или да идентифицират проблеми с видимостта в публикувано съдържание с цел преглед и организиране на образователното съдържание.

За такива бяха използвани клъстерните техники (K-Means и самоорганизиращи се карти-SOM) с цел намиране на групи със специфично поведение и откриване на лица с нетипично поведение [45, 68]. Авторът предполага, че може да е интересно да се свърже поведението по подразбиране с резултатите от оценката на учениците.

Някои автори [46] имат за цел да подобрят използването на образователните медии и да оценят ефективността на действията на възпитателя. За това авторите се използват клъстериране, правила за асоцииране и техники за дърво за вземане на решения за извличане и комбиниране на правила, свързани с образователни медии и учебни планове, с цел подпомагане на образователното управление на администратора и за подобряване на преподаването и ученето [43, 12, 71].

1.2. Оценяване

При електронното обучение оценката на предложените дейности и разбирането на моделите на поведение на учениците за преподавателите са важни и сложни въпроси и редица изследвания се стремят да намерят решения на тези въпроси [53].

В [91] се стремят да оценят дейностите на учениците в опит да идентифицират нетипични модели на поведение от дневници за достъп и дейности, реализирани от обучаемите. Авторите се стремят да извлекат полезни модели чрез статистически анализ и прилагане на обобщаваща техника от уеб-дневници, за да намерят тенденции в тези модели. Целта е да се помогне на преподавателя в оценката на обучението и да се интерпретират нуждите на учениците, за да се подобри тяхното представяне [12].

В допълнение към намаляването на чувството за изолация от студента в курса за електронно обучение, авторите се стремят и да намалят процента на отпадане на студента, използвайки техники за наблюдение и постоянно приспособяване на преподаването на студента по време на обучението [77].

В [52] се използват генерираните данни от системите за управление на обучението, за да идентифицират важни модели в отношенията между профила на ученика и решаването на проблеми, за да се установят най-добрите стратегии за разрешаване, които могат да бъдат препоръчани за различни профили на ученици. Авторите използват правила за асоцииране за откриването на тези модели [52]). Целта е да се дадат предложения на преподавателите за дидактическо-педагогически интервенции, които имат за цел подобряване на представянето на учениците по време на обучението и обучението въз основа на решаване на проблеми [59].

Някои изследователи се стремят да оценят резултатите от обучението на учениците, използвайки информация от взаимодействие и достъп до образование медии [15]. Целта е да се осигури по-добра представа за преподавателите относно основните фактори, които влияят върху учебните постижения на учениците. За постигането на тези цели се прилага

релационен анализ на Grey [14], техника на клъстериране, правила за размита асоциация и алгоритми за размито извеждане.

Тъй като няма окончателни механизми за оценка на ефективността на процеса на електронно обучение, (обикновено крайният резултат се дава чрез обобщена оценка), [12, 54] представят уеб система за реализиране на формативна оценка, използвайки статистически корелационен анализ, Fuzzy клъстеризация, наред с други техники [47, 59, 62]. Целта е да се идентифицират основните формативни правила за оценяване според профила на ученика, целящ персонализиране в използваните образователни медии. Тази система позволява на преподавателите и преподавателите да разберат факторите, които влияят върху ученето на учениците, и осигурява по-добро производство на учебни материали и най-добрите практики на преподаване [13, 54].

В [80] разработиха предсказващ модел на представянето на ученика, който въз основа на постигнатото преди това тестове за умения, е в състояние да предскаже представянето на ученика по време на учебния процес. По този начин, моделът за прогнозиране открива учене на недостатъци периодично, позволявайки на преподавателя да проследява представянето на ученика и в крайна сметка да направи подходящата интервенция, за да се предотврати избягването или неуспехът на ученика [13, 54].

Прогнозирането на поведението на ученика ви позволява да разберете аспектите, които влияят положително на обучението. Авторите [17] се стремят да получат лична информация и информация за поведението на ученика и представянето му по време на обучението. Авторите предлагат модел за самонастройка, способен да идентифицира навици, нужди и предпочитания.

От друга страна, в [79] твърдят, че наблюдението на ученето на учениците означава разбиране на тяхното поведение, нужди и недостатъци и да се установят стратегии за насочване на педагога при вземане на решения в помощ на ученика. Авторите предлагат правила за асоцииране на данни за приложението за оценка на преподавателите чрез алгоритъм Predictive-Apriori [71, 4]. Тези стратегии осигуряват по-надеждни параметри, за да помогнат на преподавателя при по-добри решения според интересите на ученика.

В [81] твърдят, че е важно да се оцени учебната система, за да се осигури динамична и подходяща ориентация към ученика, като изпълнение на дейности и за предоставяне на обратна връзка на преподавателя за поведението и представянето на ученика по време на учебния процес.

В [92] се разработи система за оценка на обучението, базирана на цифрови игри, търсейки информация за важността и качество на преподавателската практика. За анализ на генерираните данни по време на учебната практика са използвани CLOSET [61] и алгоритми за правила за асоцииране (Apriori). Подобреният CLOSET се оказва повече подходящ от традиционния алгоритъм на Apriori, тъй като подобреният CLOSET намери правила без съкращения и без нарушаване на пълнотата при извличане на модели. Резултатите показаха ефективност при оценка на качеството на образованието чрез полезна информация за профила на възпитателя и, получени важни правила за насочване на възпитателя и образованието управление по отношение на представянето на ученика [71, 79, 92].

И накрая, според [26], за да разберете важна информация за представянето на ученика, е необходимо да анализирате данни като: достъп до системата за запис (дневници); взаимодействия ученик-възпитател; и семантични данни от дискуссионни форуми и чатове. Документът се стреми да постигне изводи във връзка с представянето на ученика, като използва многослойна класификация на Perceptron и случайни горски техники [34]. Резултатите подчертават важността на извършването на изводи в ученическия индивидуално представяне и за подпомагане на педагогическите стратегии, предложени за намаляване на избягването и неуспеха на учениците [12, 13, 19].

1.3. Комуникационни канали

Различават се два вида комуникация: синхронна и асинхронна. Най-важната услуга за асинхронна комуникация е електронната поща. Платформата трябва да поддържа и групова комуникация, предлагаща услуги като дискуссионни форуми и споделено пространство, където студентите могат да обменят документи. Освен асинхронна комуникация, може би на второ място по важност, е синхронната комуникация като споделено работно място (онлайн демонстрации) и/или аудио/видео конференции.

Мненията на учениците в текстови блокове и дискуссионни форуми се анализират от [35], които се стремят да идентифицират дали чувството на ученика е положително или отрицателно. В тази статия се реализира класификацията на чувствата и мненията с помощта на Модел на Марков (НММ: Hidden Markov's Models) и методът SVM (Support Vector Machine). Този подход осигурява параметри за най-добро вземане на решения, като по този начин позволяват да се подобри качеството на електронното обучение [35, 77].

От друга страна, в [73] имат за цел да намалят натоварването на преподавател и преподавател чрез генериране на резултата от участниците (учениците) в дискуссионните форуми. Авторите използват индикаторите за обучение Learning Vectors (LV) [35] като инструмент за оценка.

В [44] предлагат съпътстваща стратегия на ученика въз основа на неговия исторически анализ и взаимодействия с преподавателя чрез чатове и дискуссионни форуми. Авторите използват техниките за клъстериране и асоцииране на историческите данни и взаимодействията на учениците и целта му е да идентифицира нивото на интерес на ученика и участието му в дейностите.

1.4. Административно осигуряване

Администраторът представлява участник, отговорен за управлението на всички етапи на преподаването и обучението. За тази роля се изискват специфични умения и знания в управлението на образованието, който трябва да помогне за еволюцията и подобряването на образователната среда [29, 67]. Повечето статии препоръчват администраторът да притежава технически умения, умения за човешко поведение (култура, убеждения, нагласи и мнения) и накрая да има общ преглед на институцията във връзка със своята структура и цели. Администраторът трябва да е запознат с въпроси като структурата на учебната среда,

качеството и наличността на учебни материали и да осигури подходящ достъп до преподавателя и ученика в системите за управление на обучението [20, 69, 82].

Административното осигуряване е най-важният аспект при системите за управление на обучението. От една страна то е свързано със сигурността на системата и осъществяване на коректен достъп до системата. От друга страна правилна оценка на обучението на ученика, от гледна точка на администратора, е да се следи взаимодействието им с учителя, за да се разбере по-добре поведението на ученика в учебната среда, за да се идентифицират неговите нужди и възможните рискове от неодобрение и избягване [67, 69].

Данните, получени чрез профила на ученика и записите за достъп в средата за електронно обучение, може да са недостатъчни за откриване на сложни действия на учениците. В статията [69] се предлага система за мониторинг на ученици и преподаватели за администраторите. За това се използват подходи, базирани на онтологии и размита логика за интерпретиране на данните. Целите на тази система са: намаляване на учебните недостатъци; подобряване на нивото на взаимодействие между ученик и възпитател; идентифициране на възможни рискове от заобикаляне; и създаване на персонализирани мерки за подобряване на представянето на учениците [68].

В [79] цитират трудностите при наблюдението на ученето на ученика. Авторите предлагат използването на правила за асоцииране за оценка на преподавателите. Целта е да се предоставят надеждни индикатори на администратора за вземане на по-добри решения в подкрепа на преподавателя в съответствие с интересите на учениците [31, 46, 69].

Статията [75] се стремят да установят индикатор за обучение в подкрепа на количествената и качествена оценка на обучението [19, 20, 29]. Този индикатор се актуализира чрез комуникация между студенти, преподаватели и администратори и е в състояние да предостави информация за резултатите на учениците за всички образователни участници в обучението и обучението. Според авторите за количествено определяне на производителността, усилията, сътрудничеството и взаимодействието могат да осигурят по-добри резултати за управлението на администраторите [28].

2. Анализ на големи масиви от данни при обучението и игрите

Анализът на големи масиви от данни при обучението (LA - learning analytics) е изследователска област, която поставя ново предизвикателство с цел да се анализира голямо количество данни, произведени от обучението и/или сериозните игрите, възможно най-ефективно и смислено. Learning analytics се отнася до машинното обучение за предоставяне на прогнози за успеха на учениците и предписанията на учащите и учителите.

По принцип повечето техники за извличане на данни са добре подходящи за LA. Проучване показва, че най-често се използват основните техники за извличане на данни от клъстеризация, правило за асоцииране, извличане на визуални данни, статистика и регресия, а някои техники, като последователно извличане на шаблони, извличане на текст, корелационно извличане, откриване на външни източници, каузално извличане и оценка на

плътност, не се използват често поради сложността в получаването на атрибутите, необходими за регулиране или адаптиране към индивидуалните нужди [27].

Във всяка образователна организация има система за управление на обучението (Learning Management Learning Systems - LMS), която поддържа основния процес на обучение. LMS създават и генерират данни от различни дейности и ресурси, като дневници на дейностите на обучаемите, дневници на дейностите на курсовете, съдържание на курса или резултати от оценките на студентите. Има спешна нужда от анализ на тези данни, за да се постигнат по-добри резултати на учениците. Синергията между Big Data, learning analytics и управлението на знанието играе нарастваща роля чрез адаптивно и персонализирано обучение, извличане на образователни данни, визуализация на данни, visual analytics и дава по-добра информираност на служителите и учители на висшето образование [49].

Moodle е един от най-широко използваните LMS. Във версията Moodle 3.4 има LA с помощта на механизми за машинно обучение, които надхвърлят обикновения дескриптивен анализ и има два вградени модела: студенти с риск от отпадане и без преподавателска дейност [50].

Това е причината, че някои изследователи разработиха специфичен инструмент за извличане на данни от Moodle и приложиха техники за клъстеринг, за да получат точните групи студенти, за да класифицират учениците [4]. Класификаторът показва какви са основните характеристики на учениците във всяка група и позволява класирането на нови онлайн студенти.

В статия [5] авторите групират студенти, като извличат данни от журнала на Moodle. Основните цели са да се определят съответните характеристики на клъстеринг и да се определи дали нашите студенти показват различно поведение на учене. Експериментът изпълни следните алгоритми за клъстериране, осигурени от Weka: Максимализация на очакванията, йерархично клъстериране, прости K-средства и X-средства. Резултатите показват, че методът е достатъчен, за да помогне за идентифициране на ученици в групи с проблеми, тъй като получаването им чрез клъстериране винаги е било добре разделено според нивото на тяхната активност, като това ниво на дейност също е свързано с оценките, които са получили в оценявани дейности.

Като цяло повечето техники за извличане на данни са много подходящи за LA. Проучване показва, че основните техники за извличане на данни на клъстериране, правило за асоцииране, визуално извличане на данни, статистика и регресия са често използвани и някои техники, като последователно извличане на шаблони, извличане на текст, корелация на корелация, откриване на извънредни данни, причинен добив и плътност оценки, не се използват често поради сложността при получаване на атрибутите, необходими за регулиране или адаптиране към индивидуалните нужди [6]. Проучването обхваща най-подходящите проучвания (402 статии), описващи приложенията на извличане на данни, свързани с обучението и LA във висшето образование, за да предостави възможности и решения на различни учебни проблеми, свързани с четири основни измерения: компютърно поддържан учебен анализ, компютърно поддържан прогнозен анализ, компютърно

поддържан поведенчески анализ и компютърно поддържан анализ на визуализация от 2000 до 2017 г. [6].

ЛА се справя с редица предизвикателства, като например справяне с нарастващ обем данни, намиране на значими показатели и подходяща визуализация на информация [2]. Изследователската работа показва колко полезно може да бъде прилагането на техники за извличане на данни в системите за управление на курсове за онлайн инструктори според техники за визуализация, за да получат общ поглед върху данните за използването на студента. Тези техники могат да се използват поотделно или да се прилагат заедно, за да се получи интересна информация по по-ефективен и по-бърз начин, като например прилагане на техники за клъстериране, за да се получат точните групи, на които учениците могат да бъдат разделени; използване на класификатор, който показва кои са основните характеристики на учениците във всяка група; и прилагане на правила за асоцииране, за да откриете дали има някаква връзка между тези характеристики и други атрибути [7].

III. Методите и техниките за препоръки и адаптация на учебно съдържание и работен поток от учебни дейности

1. Методите и техниките за адаптация на учебно съдържание и работен поток от учебни дейности

Дефиницията за адаптивни и интелигентни решения включва различни начини за добавяне на адаптивна или интелигентна функционалност към образователните системи. По принцип, всяка технология може да се реализира чрез други техники и методи, които отговарят на нейната функционалност и различни начини за прилагане (Brusilovsky, 1996). Според Брусиловски (Brusilovsky, 1997) в адаптивните хипермедийни системи може да се адаптират две неща – съдържанието на учебните страници и хипервръзките между тях, т.е. има две нива на адаптация – на ниво съдържание и на ниво връзки. Тези нива са два различни начина на хипермедийна адаптация и се наричат съответно адаптивна презентация и адаптивна навигация (Bontchev&Vassileva, 2006).

Други автори, като Сантос (Santos et al., 03) се фокусират върху адаптацията в зависимост от стила на учене, представянето на съдържание и взаимодействието или върху динамичното съставяне на учебен курс и избор на съдържание. Адаптивната презентация и адаптивната навигация в голяма част от случаите могат да бъдат декомпозирани на някои техники, споменати по-горе, отговарящи на категоризацията, въведена от Brusilovsky (Brusilovsky, '01; Brusilovsky, 1996).

Съществуват и технологиите на интелигентни системи за обучение, някои от които се припокриват отчасти или изцяло с тези от адаптивните хипермедийни системи, като подходът за съставяне на учебен план (curriculum sequencing), който принадлежи както на И на интелигентни системи за обучение, така и на адаптивните хипермедийни системи и е вариант на адаптивната навигация.

1.1 *Интелигентен анализ на решения*

Тази технология се занимава с анализиране на крайните отговори или решения на обучаемите, без значение как са получени тези отговори. За разлика от неинтелигентното проверяване на въпроси, с отбелязване само на верен или грешен отговор, интелигентните анализатори могат точно да кажат какво е грешно или непълно и коя липсваща или некоректна част от знанията е отговорна за това. Те изпращат отговори на обучаемите от направения анализ (Brusilovsky&Peulo, 2003).

1.2 *Интелигентни Технологии*

Към класическите технологии на адаптивните системи за електронно обучение, могат да се добавят някои техники от системите за извличане на информация - веб-базирано обучение и компютърно-подпомогнато съвместно обучение, които са разделени в три основни групи:

Адаптивно Филтриране на Информацията (АФИ)

АФИ представлява класическа технология от сферата на системите за извличане на информация. Нейната цел е намирането на документи, подходящи за конкретен обучаем,

спрямо зададените от него интереси, измежду множеството текст-базирани документи. Тази техника се прилага в уеб, чрез филтриране на резултатите от търсенето, подреждането им в подходящ ред и задаване на препоръки за тези документи, включени в резултатите, които са най-важни измежду всички намерени (Brusilovsky&Peylo, 2003). Въпреки, че машините на АХС, използвани от системите на АФИ, са много различни от машините на АХС на ниво интерфейс, първите много често използват адаптивна навигация.

Интелигентен контрол

Интелигентен контрол - тази технология е от областта на уеб-базираното обучение. При него учителят/инструктор не може да види признаците на разбиране и/или объркване по лицата на обучаемите. Така става трудно да се идентифицират проблемните обучаеми, които се нуждаят от допълнителни обяснения и помощ, по-бедните такива, които се нуждаят от по-задълбочени материали, както и части от учебните материали, които са прекалено лесни или прекалено трудни или неясни и объркани. С тази техника системите могат да следят всички действия на обучаемите. За да се подпомогне големия обем от данни, събирани в процеса, се използва система за Изкуствен Интелект (ИИ/АИ), подпомагащ учителите (Brusilovsky&Peylo, 2003).

Интелигентна поддръжка на сътрудничество

Това е интересна група от технологии, разработени от две области, първоначално доста отдалечени една от друга: компютърно-подпомогнатото съвместно обучение и системите за информационни технологии (Brusilovsky&Peylo, 2003). Има три основни групи технологии от този вид:

- Адаптивно формиране на група и партньорска помощ – тези технологии се занимават с формирането на групи от обучаеми за различни видове съвместни задачи (Norpe, 1995); (Ikeda et al., 1997)
- Адаптивна поддръжка на сътрудничество – тази група от технологии се опитва да осигури интерактивна поддръжка на процеса на сътрудничество. Използвайки шаблони за „добри“ и „лоши“ сътрудничества, системи като COLER (Gonzalez, 2003) или EPSILON (Soller&Lesgold, 2003), могат да се обучават и съветват участниците от един екип как да си сътрудничат;
- Виртуални обучаеми - сравнително стара технология, която се опитва да въведе различни видове виртуални участници в една учебна среда (Frasson et al., 1996)

1.3 *Техники за адаптация на учебно съдържание и работен поток от учебни дейности*

Основните техники за адаптация на учебно съдържание и работен поток от учебни дейности се разделят на две големи групи: адаптивна презентация и адаптивна навигация. Основната идея на адаптивната презентация е да се адаптира съдържанието на поискана от даден потребител страница към неговия профил, който съдържа настоящите му познания, цели и други негови характеристики. Тази техника се дели на два основни метода: адаптивна презентация на текст и адаптивна презентация на модалност.

Под адаптивна презентация на текст се разбира представяне на различно текстово съдържание на една и съща страница за различните потребители в различни периоди от време. Основната цел е да се адаптира съдържанието на една хипермедийна страница в съответствие с потребителските цели, знания, умения за работа с медии и всяка информация, която се съхранява в потребителския модел. Тази адаптация има два под-вида – адаптивно представяне на фрагмент и адаптация, базирана на естествения език (Brusilovsky, 2001), като се подразбира, че втория вид използва първия, но с използване на технологии, базирани на естествения език.

Целта на адаптивната навигация е да помага на обучаемия да намери своя най-подходящ път в хипер-пространството, чрез адаптивно представяне на връзките в него, съобразено с неговия ПМ. Тя помага на обучаемия да избере как да премине от текущата страница към следващата, в зависимост от неговите нужди, ниво на познание, интереси и др.

Основните техники за адаптация на учебно съдържание и работен поток от учебни дейности са обобщени в табл. 1.

Таблица 1. Основните техники за адаптация на учебно съдържание и работен поток от учебни дейности

Адаптивна Презентация	Адаптивна презентация на текст	Асоцииран с условия текст (Conditional Text)	Страницата се разделя на части. Всяка част е свързана с едно или повече условия за ниво на потребителски знания. Съдържанието на цялата страница се показва на потребителя, само ако той отговаря на съответните условия
		Разпъване на текст (Stretch text)	Във всяка стандартна хипертекстова страница всяка връзка е свързана с друга отделна страница или ресурс. В тази техника хипертекстовите връзки са асоциирани с текст. Ако обучаемият желае, той може да види избрания текст, ако избере съответната връзка или да бъде отново скрит за да бъде по-компактно изображението.
		Сортиране на текст	Техника пренарежда съдържанието в дадена страница. Вариантите за пренареждане трябва да бъдат предефинирани и след това да се направи избор, в зависимост от потребителския модел.
		Добавяне/премахване на фрагменти (Inserting/Removing Fragment)	Базирана на фреймове техника, която е разширение на първата от споменатите (асоцииран с условия текст). Съдържанието на всяка една концепция е представено в различни празни фреймове, разделени на части. Тези части съдържат обяснения, връзки с други фреймове, примери и др. Използват се правила, които определят кои части да бъдат показани на даден обучаем и в какъв ред.

		Блед текст (Text Fragment Dimming)	Някои части от текста, които не се считат за важни, се показват с форма, която е по-малко видима (по-малки букви или с по-нисък контраст). В някои случаи това е по-добре, отколкото текста да бъде скрит, но е важен за разбирането на учебния материал или обучаемият вече да е запознат с него, но е добре да се припомни.
		Варианти на страница	Вместо да се показват адаптации на страницата, може да се показват предварително създадени страници, като варианти за едно и също обяснение и да се показва само една от тях, в зависимост от неговия потребителски модел
	Адаптивна презентация на модалност		Съществува голям набор и избор от различни видове медии, които да представят на обучаемия като текст, видео, музика, анимация и др. Фрагменти от различни медии могат да представят едно и също съдържание и така системата може да избере най-подходящата за даден обучаем за даден момент.
Адаптивна навигация	Предефинирани връзки (direct guidance)		Може да бъде прилагана във всяка система, която може да реши кой е следващият „най-добър“ възел за обучаемия, според дадени параметрите.
	Сортиране на връзки		Тази технология сортира всички връзки в една страница в зависимост от дадени критерии, като колкото по-напред е една връзка, толкова е по-подходяща за даден обучаем.
	Скриване на връзки		Скриват се връзките към неподходящите страници, като така се предпазва потребителя от излишна информация и ограничава комплексността на хиперпространството
	Анотация на връзки		Този метод дава възможност да бъде описана страницата или ресурса, към който връзките водят.
	Последователност на учебния план		Целта е да се осигури учебен план, който да включва подходяща за всеки обучаем последователност от учебни единици, които той трябва да усвои, както и последователност от учебни задачи, които той трябва да реши. Това работи на принцип, подобен на предефинирани връзки (direct guidance), но гарантира дългосрочна последователност, без ограничаване на хиперпространството.

	Адаптивни навигационни карти (adaptive navigation maps)	Отнася се за личния преглед на хиперпространството. Обикновено се показват само връзките, които се считат за подходящи.
	Адаптивно генериране на връзки (adaptive link generation)	Генерира динамичен списък с допълнителни подходящи връзки. За тази цел използва подходи като откриване на връзки за навигация, базирана на сходство между елементите и др.

2. Техники за препоръка на учебно съдържание. Категории системи за препоръки.

Системите за препоръчване са подклас на системите за филтриране на информация, които имат основна цел да предвидят рейтинг или теми за предпочитане, които потребителят би дал на даден елемент. Системата изчислява и му предоставя подходящо съдържание въз основа на познанията за самия елемент, съдържанието и взаимодействията между потребителя и артикула (елемента). Тези системи обикновено имат широко приложение в популярните търговски приложения.

Основните области, в които намират приложение подобни системи са: генератори на плейлисти за видео и музикални услуги като Youtube, Netflix и Spotify, препоръчители на продукти и услуги като Amazon, или препоръчители на съдържание в социални платформи като Facebook и Twitter. Подобни системи се използват активно и за препоръчване на специфични теми като ресторанти и онлайн запознанства.

2.1. Основни цели на системата за препоръка

Системите с вградена функционалност за препоръки чрез машинно обучение, придобиват все по-голямо значение през деветдесетте години, откакто интернет се превърна във важна среда за бизнес и транзакции за електронна търговия. Това е така, защото глобалната мрежа предоставя безпрецедентни възможности за персонализация, които не са налични до въпросния момент от време [93]. Сърфирането в интернет прави лесно събирането на данни за поведението на потребителите. Основен принцип при препоръките е, че съществуват значителни зависимости между потребител и артикул - под артикул, следва да се разбира стока, услуга и т.н., а в рамките на текущото изследване артикул, следва да се разбира учебен ресурс. В много случаи има зависимост и корелации между цели категории артикули. В други случаи зависимостите могат да бъдат просто между отделни елементи, а не категории. Всички тези особености могат да бъдат научени по начин, фокусиран върху данни от матрица за оценяване и полученият модел, да се използва за прогнозиране на интересите на целевите потребител. Разбира се, колкото по-голям е броят на оценяваните елементи, както и колкото по-голяма е активността на даден потребител, толкова по-лесно ще е да се правят акуратни прогнози за неговите предпочитания. Основните цели пред всяка една система за препоръчване са:

- Релевантност;

- Новост;
- Случайност;
- Разнообразие.

Релевантност

Най-очевидната оперативна цел на системата за препоръчване е да препоръчва елементи, които са наистина от съществено значение за потребителя. Потребителите консумират (купуват или гледат) артикули, в това число да се включват и събития, които им се струват интересни. Да бъдат *релевантни* е основната оперативна цел на една система за препоръки. Въпреки това това е необходимо, но недостатъчно условие за качествени препоръки. Затова по-долу следват някои вторични цели, които допринасят за високото качество на опериране на системата за препоръки.

Новост

Системите за препоръчване са наистина полезни, когато препоръчаният елемент е нещо, което потребителят не е виждал в миналото. Нови препоръки са препоръки за артикули, за които потребителят не знае. Новостта за даден потребител *u* се дефинира от съотношението на неизвестните елементи в списъка на топ *N* препоръчани елементи [94]:

$$\text{Новост}(u) = \sum_{t \in R} (1 - f(u, i)) |R| \quad (1)$$

където *R* е множество от *N* препоръчани елементи, *f(u, i)* е двоична функция, която връща 1 за резултат, когато потребителят вече знае за елемент *i*, и 0, когато потребителят не е запознат с *i*.

Важно е да се направи уточнение, че системата за препоръчване не би могла еднозначно, т.е. точно, да прецени дали даден потребител знае за даден артикул: събитие, услуга и т.н. Затова е разумно допускането, че колкото по-малко популярни или непопулярни са препоръчаните артикули, толкова по-голям е шансът те да са нови за потребителя.

Случайност

Визира се, че препоръчваните елементи са някак неочаквани и случайни, с лек елемент на късмет и изненада [93]. Тук е важно да се отбележи, че става въпрос, за това, че препоръките са били наистина изненадващи за потребителя, а не просто нещо, за което не са знаели преди това. Често би могло да се случи така, че определен потребител да консумира елементи от определен тип, а да има латентен интерес към други категории елементи, за които дори самият потребител следва да намери за изненадващи. Например, ако в един квартал се отвори нов индийски ресторант, препоръката за него към потребител, който обикновено яде индийска храна, би била нова, но не непременно случайна. От друга страна, когато на този потребител се препоръча етиопска храна, като за него не е известно, че тази храна може да му хареса, то тогава такава препоръка се нарича случайна. Случайността при нея има и благоприятен страничен ефект в увеличаване на разнообразието на препоръчаните артикули, което е предпоставка за нова тенденция на интерес у

потребителя. Разбира се, тук съществува и известен риск от предлагане на неподходящи артикули, т.е. такива, които не биха се харесали на потребителя от страна на системата.

2.2. Категории системи за препоръки. Основни характеристики.

Системи за препоръки със съвместно (Collaborative) филтриране

Моделите на системи със съвместно филтриране, използват силата на оценките (рейтинги), предоставени от множество потребители, за да отправят препоръки. Основната идея на този тип методи е, че неопределените рейтинги могат да бъдат „преписани“, защото оценките които потребителите дават, често са силно свързани [95].

Съществуват няколко основни подхода при съвместното филтриране – базирани на паметта (memory based) и базирани на модел [93]. Базираните на паметта или още познати като базирани на съседство (neighborhood based) се делят още на:

1) Потребителско базирано съвместно филтриране. Тук оценките, предоставени от сходни потребители на потребител А, се използват, за да се отправят препоръки за А. Основната идея е да се определят група потребители, които са „подобни“ (или с подобен вкус) на потребител А, като по този начин да прогнозираят оценки за неопределените артикули от А, като се изчисляват средни стойности на рейтингите на тази група потребители.

2) Съвместно филтриране, базирано на артикули. За да се направят прогнози за рейтинг за целеви елемент В от потребител А, първата стъпка е да се определи набор от S елементи, които са най-сходни с целевия елемент В. Оценките в набора от елементи S, които са посочени от А, се използват за прогнозиране дали на потребителя А ще се хареса елемент В.

Базираните на модел методи се използват в контекста на прогнозните модели. В случаите, когато моделът е параметризиран, параметрите на този модел се научават в контекста на рамка за оптимизация - optimization framework. Някои примери за такива базирани на модели методи включват дървета на решения, модели, базирани на правила методи, методи на Бейс и модели на латентни фактори. Много от тези методи като модели с латентни фактори имат високо ниво на покритие дори за много оскъдни матрици.

Системи за препоръки базирани на съдържание

В тези системи обектите се определят главно от свързаните с тях характеристики. Системата научава профила на интересите на новия потребител въз основа на присъстващите функции в обекти, които потребителят е оценил. Артикулите в тази система биват разглеждани в контекста на ключовите думи, които ги описват. В подобни системи, за да се направят препоръки се използват описателните атрибути (ключови думи) на артикулите. Терминът „съдържание“ се отнася именно до тези описателни атрибути или

характеристики. При методите, базирани на съдържание, оценките и поведението на потребителите се комбинират с информацията за съдържанието налична в артикулите.

В системите базирани на съдържание, описанието на даден артикул, който е етикетирани (labeled) с рейтинг, се използва като тренировъчни данни за създаване на специфичен за потребителя класификационен проблем - или още познато като регресионно-моделиран проблем. Класовата или зависимата променлива съответства на специфични оценки за даден артикул. Тези тренировъчни документи се използват за създаване на класификационен или регресионен модел, който е специфичен за активния потребител. Този модел се използва за предвиждане на неизвестна оценка за съответен артикул. Тези системи имат известно предимство пред съвместните системи за предлагане за нови артикули, когато няма достатъчно рейтинги за този артикул [3]. Това е така, защото други артикули с подобни характеристики или съдържание, например ключови думи, е възможно да бъдат оценени от същия потребител. Така системата би могла да използва оценки във връзка с атрибутите на артикула, за да дава препоръки, дори когато не съществуват рейтинги за този артикул.

Системите за препоръка, базирани на съдържание, имат и своите слаби места. Например в много от случаите дават очевидни препоръки именно поради използването на ключови думи. Така ако даден потребител никога не е консумирал даден артикул с определен набор от ключови думи, то същият или подобен артикул (със тези ключови думи) няма шанс да бъде препоръчан. Също така въпреки, че се спомена, че тези системи са добри при предоставянето на препоръки за нови артикули, те обаче не са така ефективни при предоставянето на препоръки за нови потребители [93].

Системи за препоръки на базата на знания

Системите за препоръки на базата на знания са особено полезни в контекста на артикули, които не се купуват много често [93]. Примерите включват вещи като недвижими имоти, автомобили, заявки за туризъм, финансови услуги или скъпи луксозни стоки. В такива случаи би могло да липсват достатъчно оценки за процеса на препоръчване, тъй като артикулите се купуват рядко и обикновено са с различни опции, и е трудно да се получи достатъчен брой оценки за конкретна инстанция на съответния артикул. Подобна ситуация съществува и при проблема със студения старт (cold-start problem), при който няма достатъчно оценки за процеса на препоръчване. Освен това потребителските предпочитания търпят развитие и изменение с течение на времето.

Друг проблем е, че даден артикул е възможно да има атрибути, които да кореспондират на конкретни негови характеристики, а потребителят да се интересува само от артикули с тези характеристики и свойства.

Такива случаи могат да бъдат решени с помощта на препоръчителни системи, базирани на знания, в които не се използват оценки за целите на препоръките. Тук по-скоро, в процеса по препоръчване, се използват сходства между изискванията на клиента и описанията на артикулите. Процесът се съпровожда от употребата на базови знания, които съдържат данни

за правила и функции за сходност, т.е. такива, които изчисляват сходност. Този процес на препоръчване се основава на тези базови знания и от там идва и неговото име. Изричната спецификация на изискванията на потребителите води до по-голям контрол на клиента над процеса на препоръчване.

]Системите за препоръки от този вид могат да бъдат разделени на няколко подтипа:

1) Препоръчителни системи на базата на конкретни случаи - в този род системи потребителят определя конкретни отправни (целови) точки. Например потребителят определя, че е заинтересован от къща с определен брой бани, обзавеждане, ценови диапазон и т.н. Тези отправни точки се изпращат в заявката към системата за препоръки, която връща резултат, който удовлетворява критериите на потребителя. Тези резултати обикновено се използват за нови целови точки, като потребителят има възможност да променя част от посочените атрибути, за да таргетира определени артикули. По този начин се сформира процес, който насочва потребителя към конкретни артикули, представляващи неговия интерес;

2) Препоръчителни системи на базата на ограничения - При тях потребителите изрично посочват долни и горни граници на конкретни изисквания и/или ограничения към атрибутите на артикула. Системата използва правила, специфични за областта си на действие, съответстващи на потребителските изисквания към атрибутите на артикула. Потребителят има възможност да регулира тежестта на всеки един от тези атрибути в процеса на търсене, като самият процес на търсене се повтаря многократно и интерактивно, подобно на предишния тип системи на базата на конкретни случаи, докато потребителят не получи желан резултат.

Таблица 2. Сравнителен анализ на типовете системи за препоръки

Подход	Цел	Входни данни
Съвместно филтриране	Предлагане на препоръки, базирани на съвместен подход, който използва оценките и действията на останалите потребители	Оценки на потребителя и оценки на останалите потребители
Базирана на съдържание	Предлагане на препоръки въз основа на съдържанието (съответни атрибути), които потребителят е предпочитал, базирано на оценки и конкретни действия	Оценки на потребителя и атрибути на артикулите
Базирано на знания	Предлагане на препоръки въз основа на изрична спецификация на типа съдържание (атрибути), които потребителят е изискал	Спецификация на потребителя, атрибути на артикулите и познания в конкретната област.

Използвана литература

- [1] Adeva, J., Carroll, N., Calvo, R., 2006. Applying plagiarism detection to engineering education. In: Information Technology Based Higher Education and Training, 2006. ITHET '06. 7th International Conference on, pp. 722–731. <http://dx.doi.org/10.1109/ITHET.2006.339692>.
- [2] Akerkar, R., Bassiliades, N., Davies, J., Ermolayev, V., 2014. Research and applications in web intelligence, mining, and semantics. In: Proceedings of the 4th International Conference on Web Intelligence, Mining and Semantics (WIMS14), number 0 in WIMS '14. ACM, New York, NY, USA pp. 0:1–0:6. ISBN 978-1-4503-2538-7.
- [3] Alhajj, R., 2007. In: Advanced Data Mining and Applications: Third International Conference, ADMA 2007, Harbin, China, August 6–8, 2007 Proceedings. vol. 4632, Springer Science & Business Media.
- [4] Anupama Kumar, S., Vijayalakshmi, M., 2012. Mining of student academic evaluation records in higher education. In: Recent Advances in Computing and Software Systems (RACSS), 2012 International Conference on, pp. 67–70. <http://dx.doi.org/10.1109/RACSS.2012.6212699>.
- [5] Baker, R.S., Yacef, K., 2009. The state of educational data mining in 2009: a review and future visions. J. Educ. Data Min. 1 (1), 3–17 ISSN 2157-2100.
- [6] Bhushan, N., Rai, K., 2004. Strategic Decision Making – Applying the Analytic Hierarchy Process. Springer-Verlag, London 978-1-85233-864-0 <http://dx.doi.org/10.1007/b97668>.
- [7] Buckley, J.J., 1985. Fuzzy hierarchical analysis. Fuzzy Sets Syst. 17 (3), 233–247. [http://dx.doi.org/10.1016/0165-0114\(85\)90090-9](http://dx.doi.org/10.1016/0165-0114(85)90090-9).
- [8] Bunkar, K., Singh, U., Pandya, B., Bunkar, R., 2012. Data mining: prediction for performance improvement of graduate students using classification. In: Wireless and Optical Communications Networks (WOCN), 2012 Ninth International Conference on, pp. 1–5. <http://dx.doi.org/10.1109/WOCN.2012.6335530>.
- [9] Caballe, S., Xhafa, F., Abraham, A., 2008. Towards an automatic real-time assessment of online discussions in computer-supported collaborative learning practices. In: Digital Information Management, 2008. ICDIM 2008. Third International Conference on, pp. 470–475. <http://dx.doi.org/10.1109/ICDIM.2008.4746758>.
- [10] Calinger, M., Howard, B.C., 2008. Evaluating educational technologies: a historical context. IJICTE 4 (4), 9–18. <http://dx.doi.org/10.4018/jicte.2008100102>.
- [11] Chaczko, Z., Dobler, H., Jacak, W., Klempous, R., Maciejewski, H., Nikodem, J., Nikodem, M., Rozenblit, J., Araujo, C., Sliwinski, P., 2006. Assessment of the quality of teaching and learning based on data driven evaluation methods. In: Information Technology Based Higher Education and Training, 2006. ITHET '06. 7th International Conference on, pp. 21–36. <http://dx.doi.org/10.1109/ITHET.2006.339723>.

- [12] Chen, C.-M., Chen, M.-C., 2009. Mobile formative assessment tool based on data mining techniques for supporting web-based learning. *Comput. Educ.* 52 (1), 256–273., <http://dx.doi.org/10.1016/j.compedu.2008.08.005>. ISSN 0360-1315.
- [13] Chen, C.-M., Chen, Y.-Y., Liu, C.-Y., 2007. Learning performance assessment approach using web-based learning portfolios for e-learning systems. *Syst. Man Cybern. Part C* 37 (6), 1349–1359. <http://dx.doi.org/10.1109/TSMCC.2007.900641>. ISSN 1094-6977.
- [14] Chen, F.-S., Chang, T.-C., Liao, H.-H., 2000. The application of the grey relation analysis on teacher appraisal. In: *Systems, Man, and Cybernetics, 2000 IEEE International Conference on*. vol. 5. pp. 3366–3371. <http://dx.doi.org/10.1109/ICSMC.2000.886525>.
- [15] Chen, S.Y., 2005. Evaluating the learning effectiveness of using web-based instruction: an individual differences approach. *IJICTE* 14. <http://dx.doi.org/10.4018/jicte.2005010107>
- [16] Chen, Y., Mohd Taib, S., Che Nordin, C., 2012. Determinants of student performance in advanced programming course. In: *Internet Technology and Secured Transactions, 2012 International Conference For*, pp. 304–307.
- [17] Da, M., Hai-guang, H., Jian-he, G., 2010. The naive bayesian approach in classifying the learner of distance education system. In: *Information Engineering and Computer Science (ICIECS), 2010 2nd International Conference on*, pp. 1–4. <http://dx.doi.org/10.1109/ICIECS.2010.5677687>.
- [18] Dejaeger, K., Goethals, F., Giangreco, A., Mola, L., Baesens, B., 2012. Gaining insight into student satisfaction using comprehensible data mining techniques. *Eur. J. Oper. Res.* 218 (2), 548–562. <http://dx.doi.org/10.1016/j.ejor.2011.11.022>. ISSN 0377-2217.
- [19] Deng, J., Hu, J., Chi, H., Wu, J., 2010. A study of teaching evaluation in adult higher education based on decision tree. In: *Information Technology and Computer Science (ITCS), 2010 Second International Conference on*, pp. 381–385. <http://dx.doi.org/10.1109/ITCS.2010.99>.
- [20] Dongsheng, Z., Wenjing, J., 2009. Design and implementation of university educational decision support system on the students satisfaction survey. In: *Computer Science-Technology and Applications, 2009. IFCSTA '09. International Forum on*. vol. 3. pp. 428–430. <http://dx.doi.org/10.1109/IFCSTA.2009.344>.
- [21] Drigas, A., Vrettaras, J., 2006. An intelligent search engine assessing learning material to improve learning procedures. In: *Information Technology Based Higher Education and Training, 2006. ITHET '06. 7th International Conference on*, pp. 875–883. <http://dx.doi.org/10.1109/ITHET.2006.339713>.
- [22] Eagle, M.J., Barnes, T., 2013. Evaluation of automatically generated hint feedback. In: *D'Mello, S.K., Calvo, R.A., Olney, A. (Eds.), Proceedings of the 6th International Conference on Educational Data Mining. International Educational Data Mining Society, International Educational Data Mining Society, Memphis, TN, USA*, pp. 372–374 ISBN 978-0-9839525-2-7.

- [23] Fernández-Luna, J.M., Huete, J.F., MacFarlane, A., Efthimiadis, E.N., 2009. Teaching and learning in information retrieval. *Inf. Retrieval* 12 (2), 201–226. <http://dx>.
- [24] Gobert, J.D., Kim, Y.J., Pedro, M.A.S., Kennedy, M., Betts, C.G., 2015. Using educational data mining to assess students' skills at designing and conducting experiments within a complex systems microworld. *Think. Skills Creativity* 18, 81–90. <http://dx.doi.org/10.1016/j.tsc.2015.04.008>. ISSN 1871–1871, 21st Century Skills: International Advancements and Recent Developments.
- [25] Gong, M., 2008. Personalized e-learning system by using intelligent algorithm. In: *Knowledge Discovery and Data Mining, 2008. WKDD 2008. First International Workshop on*, pp. 400–401. <http://dx.doi.org/10.1109/WKDD.2008.152>.
- [26] Guleria, P., Sood, M., 2014. Data mining in education: a review on the knowledge discovery perspective. *Int. J. Data Min. Knowledge Manage. Process* 4 (5), 47. <http://dx.doi.org/10.5121/ijdkp.2014.4504>. ISSN 2230-9608.
- [27] H. Aldowah, H. Al-Samarraie, and W. M. Fauzy, “Educational data mining and learning analytics for 21st century higher education: A review and synthesis”, *Telematics and Informatics*, vol. 37. Elsevier Ltd, pp. 13–49, 01-Apr-2019, doi: 10.1016/j.tele.2019.01.007
- [28] Hien, N.T.N., Haddawy, P., 2007. A decision support system for evaluating international student applications. In: *Frontiers In Education Conference - Global Engineering: Knowledge Without Borders, Opportunities Without Passports, 2007. FIE '07. 37th Annual*, <http://dx.doi.org/10.1109/FIE.2007.4417958>. pp. F2A-1-F2A-6.
- [29] Hongxia, J., Yao, H., 2008. Classroom teaching quality evaluation based on neuro-fuzzy id3 algorithm. In: *Computational Intelligence and Design, 2008. ISCID '08. International Symposium on*. vol. 1. pp. 166–169. <http://dx.doi.org/10.1109/ISCID.2008.127>.
- [30] Howard, S.K., Ma, J., Yang, J., 2016. Student rules: exploring patterns of students' computer-efficacy and engagement with digital technologies in learning. *Comput. Educ.* 101, 29–42. <http://dx.doi.org/10.1016/j.compedu.2016.05.008>. ISSN 0360-1315.
- [31] Huebner, R.A., 2013. A survey of educational data-mining research. *Res. Higher Educ. J.* 19 (4) ISSN 1941-3432.
- [32] Ivancevic, V., Celikovic, M., Lukovic, I., 2012. The individual stability of student spatial deployment and its implications. In: *Computers in Education (SIIE), 2012 International Symposium on*, pp. 1–4.
- [33] Jindal, R., Borah, M.D., June 2013. A survey on educational data mining and research trends. *Int. J. Database Manage. Syst.* 5 (3), 53–73.
- [34] Kaur, P., Singh, M., Josan, G.S., 2015. Classification and prediction based data mining algorithms to predict slow learners in education sector. *Proc. Comput. Sci.* 57, 500–508. <http://dx.doi.org/10.1016/j.procs.2015.07.372>. ISSN 1877-0509. 3rd International Conference on Recent Trends in Computing 2015 (ICRTC-2015).

- [35] Kechaou, Z., Ben Ammar, M., Alimi, A., 2011. Improving e-learning with sentiment analysis of users' opinions. In: Global Engineering Education Conference (EDUCON), 2011 IEEE, pp. 1032–1038. <http://dx.doi.org/10.1109/EDUCON.2011.5773275>.
- [36] Kitchenham, B., Pearl Brereton, O., Budgen, D., Turner, M., Bailey, J., Linkman, S., 2009. Systematic literature reviews in software engineering – a systematic literature review. *Inf. Softw. Technol.* 51 (1), 7–15. <http://dx.doi.org/10.1016/j.infsof.2008.09.009>. ISSN 0950-5849. Special Section – Most Cited Articles in 2002 and Regular Research Papers.
- [37] Kohonen, T., 1982. Self-organized formation of topologically correct feature maps. *Biol. Cybern.* 43 (1), 59–69. <http://dx.doi.org/10.1007/BF00337288>. ISSN 1432-
- [38] Kohonen, T., 1995. *Learning Vector Quantization*. Springer, Berlin, Heidelberg 978-3-642-97610-0 <http://dx.doi.org/10.1007/978-3-642-97610-06>.
- [39] Kwan Lau, I., Fong, J., 2003. Investigation on the effectiveness on web-based learning using web-mining approach. In: *Database and Expert Systems Applications, 2003. Proceedings. 14th International Workshop on*, pp. 302–316. <http://dx.doi.org/10.1109/DEXA.2003.1232040>.
- [40] Lajis, A., Aziz, N., 2010. NI scoring technique for the assessment of learners' understanding. In: *Computer Research and Development, 2010 Second International Conference on*, pp. 379–383. <http://dx.doi.org/10.1109/ICCRD.2010.68>.
- [41] Li, R.-J., 1999. Fuzzy method in group decision making. *Comput. Math. Appl.* 38 (1), 91–101. [http://dx.doi.org/10.1016/S0898-1221\(99\)00172-8](http://dx.doi.org/10.1016/S0898-1221(99)00172-8). ISSN 0898-1221.
- [42] Liu, C., Chen, H., Tan, Y., Wu, L., 2010. The design of e-learning system based on semantic wiki and multi-agent. In: *Education Technology and Computer Science (ETCS), 2010 Second International Workshop on*. vol. 3. pp. 495–498. <http://dx.doi.org/10.1109/ETCS.2010.364>.
- [43] Liu, F.-J., Shih, B.-J., 2007. Learning activity-based e-learning material recommendation system. In: *Multimedia Workshops, 2007. ISMW '07. Ninth IEEE International Symposium on*, pp. 343–348. <http://dx.doi.org/10.1109/ISM.Workshops.2007.64>.
- [44] Liu, H., Xia, Y., 2011. Teaching evaluation system based on association rule mining. In: *Circuits, Communications and System (PACCS), 2011 Third Pacific-Asia Conference on*, pp. 1–3. <http://dx.doi.org/10.1109/PACCS.2011.5990335>.
- [45] Liu, L., 2005. Evaluating online learning applications: development of quality-related models. *Int. J. Inf. Commun. Technol. Educ.* 1 (3), 21–35. <http://dx.doi.org/10.4018/jicte.2005070102>.
- [46] Liu, Q., Peng, Y., 2013. A method of unstructured information process in computer teaching evaluation system based on data mining technology. In: *Communication Systems and Network Technologies (CSNT), 2013 International Conference on*, pp. 688–692. <http://dx.doi.org/10.1109/CSNT.2013.147>.

- [47] "London, A., Pelyhe, A., Holló, C., Németh, T., 2015. Applying graph-based data mining concepts to the educational sphere. In: Proceedings of the 16th International Conference on Computer Systems and Technologies, CompSysTech '15. ACM, New York, NY, USA, pp. 358–365. <http://dx.doi.org/10.1145/2812428.2812436>.
- [48] ISBN 978-1-4503-3357-3."
- [49] Machado, L.D.S., Becker, K., 2003. Distance education: a web usage mining case study for the evaluation of learning sites. In: Advanced Learning Technologies, 2003. Proceedings. The 3rd IEEE International Conference on, pp. 360–361. <http://dx.doi.org/10.1109/ICALT.2003.1215123>.
- [50] J. Liebowitz, "Thoughts on recent trends and future research perspectives in big data and analytics in higher education," in Big Data and Learning Analytics in Higher Education: Current Theory and Practice, Springer International Publishing, 2016, pp. 7–17.
- [51] Moodle, "Analytics - MoodleDocs," 2019. [Online]. Available: <https://docs.moodle.org/36/en/Analytics>. [Accessed: 13-Mar-2020].
- [52] Marbouti, F., Diefes-Dux, H.A., Madhavan, K., 2016. Models for early prediction of at-risk students in a course using standards-based grading. *Comput. Educ.* 103, 1–15. <http://dx.doi.org/10.1016/j.compedu.2016.09.005>. ISSN 0360-1315.
- [53] Minaei-Bidgoli, B., Kortemeyer, G., Punch, W., 2004. Association analysis for an online education system. In: Information Reuse and Integration, 2004. IRI 2004. Proceedings of the 2004 IEEE International Conference on, pp. 504–509. <http://dx.doi.org/10.1109/IRI.2004.1431511>.
- [54] Mohamad, S.K., Tasir, Z., 2013. Educational data mining: a review. *The 9th International Conference on Cognitive Science* 97, 320–324. <http://dx.doi.org/10.1016/j.sbspro.2013.10.240>. ISSN 1877-0428.
- [55] Mylonas, P., Tzouveli, P., Kollias, S., 2004. Towards a personalized e-learning scheme for teachers. In: Advanced Learning Technologies, 2004. Proceedings. IEEE International Conference on, pp. 560–564. <http://dx.doi.org/10.1109/ICALT.2004.1357477>.
- [56] Nižnan, J., Pelánek, R., Řihák, J., 2015. Student models for prior knowledge estimation. In: Santos, O.C., Boticario, J.G., Romero, C., Pechenizkiy, M., Merceron, A., Mitros, P., Luna, J.M., Mihaescu, C., Moreno, P., HersHKovitz, A., Ventura, S., Desmarais, M. (Eds.), Proceedings of the 8th International Conference on Educational Data Mining, pp. 109–116 Madrid, Spain. International Conference on Educational Data Mining (EDM) 2015, International Educational Data Mining Society (IEDMS). ISBN 978-84-606-9425-0.
- [57] Nodenot, T., Loustau, P., Gaio, M., Sallaberry, C., Lopisteguy, P., 2006. From electronic documents to problem-based learning environments: an ongoing challenge for educational modeling languages. In: Information Technology Based Higher Education and Training, 2006. ITHET '06. 7th International Conference on, pp. 280–291. <http://dx.doi.org/10.1109/ITHET.2006.339776>.

- [58] Otsuka, J., da Rocha, H., Beder, D., 2007. A multi-agent formative assessment support model for learning management systems. In: *Advanced Learning Technologies, 2007. ICAALT 2007. Seventh IEEE International Conference on*, pp. 85–89. <http://dx.doi.org/10.1109/ICALT.2007.21>.
- [59] Pang, B., Lee, L., Vaithyanathan, S., 2002. Thumbs up?: Sentiment classification using machine learning techniques. In: *Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing – Volume 10, EMNLP '02. Association for Computational Linguistics, Stroudsburg, PA, USA*, pp. 79–86. <http://dx.doi.org/10.3115/1118693.1118704>.
- [60] Pascual-Cid, V., Vigentini, L., Quixal, M., 2010. Visualising virtual learning environments: case studies of the website exploration tool. In: *Information Visualisation (IV), 2010 14th International Conference*, pp. 149–155. <http://dx.doi.org/10.1109/IV.2010.31>.
- [61] Pedro, M.O.S., Snow, E., Baker, R., McNamara, D., Heffernan, N., 2015. Exploring dynamical assessments of affect, behavior, and cognition and math state test achievement. In: Santos, O.C., Boticario, J.G., Romero, C., Pechenizkiy, M., Merceron, A., Mitros, P., Luna, J.M., Mihaescu, C., Moreno, P., Hershkovitz, A., Ventura, S., Desmarais, M. (Eds.), *Proceedings of the 8th International Conference on Educational Data Mining. International Conference on Educational Data Mining (EDM) 2015, International Educational Data Mining Society (IEDMS), Madrid, Spain*, pp. 85–92 ISBN 978-84-606-9425-0.
- [62] Pei, J., Han, J., Mao, R., 2000. Closet: an efficient algorithm for mining frequent closed itemsets. In: *Intelligent Database Systems Research Lab*, pp. 21–30.
- [63] Peña-Ayala, A., 2014. Educational data mining: a survey and a data mining-based analysis of recent works. *Expert Syst. Appl.* 41 (4), 1432–1462. <http://dx.doi.org/10.1016/j.eswa.2013.08.042>. ISSN 0957-4174.
- [64] Qingxian, P., Linjie, Q., Lanfang, L., 2009. Data mining and application of teaching evaluation based on association rules. In: *Computer Science Education, 2009. ICCSE '09. 4th International Conference on*, pp. 1404–1407. <http://dx.doi.org/10.1109/ICCSE.2009.5228194>.
- [65] Qiu, R.G., Lee, D., 2013. Transformative education web 2.0 systems for enriching high school stem education. In: *Software Engineering and Service Science (ICSESS), 2013 4th IEEE International Conference on*, pp. 352–356. <http://dx.doi.org/10.1109/ICSESS.2013.6615322>.
- [66] Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., Riedl, J., 1994. Grouplens: an open architecture for collaborative filtering of netnews. In: *Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work, CSCW '94. ACM, New York, NY, USA*, pp. 175–186. <http://dx.doi.org/10.1145/192844.192905>. ISBN 0-89791-689-1.

- [67] Robles, G., Gonzalez-Barahona, J., 2013. Mining student repositories to gain learning analytics. an experience report. In: Global Engineering Education Conference (EDUCON), 2013 IEEE, pp. 1249–1254. <http://dx.doi.org/10.1109/EduCon.2013.6530267>.
- [68] Romero, C., Ventura, S., 2007. Educational data mining: a survey from 1995 to 2005. *Expert Syst. Appl.* 33 (1), 135–146. <http://dx.doi.org/10.1016/j.eswa.2006.04.005>. ISSN 0957-4174.
- [69] Romero, C., Ventura, S., 2010. Educational data mining: a review of the state of the art. *IEEE Trans. Syst. Man Cybern. Part C* 40 (6), 601–618. <http://dx.doi.org/10.1109/TSMCC.2010.2053532>. ISSN 1094-6977.
- [70] Rosales, G., Borges de Araujo, R., Otsuka, J., da Rocha, R., 2011. Using logical sensors network to the accurate monitoring of the learning process in distance education courses. In: *Advanced Learning Technologies (ICALT), 2011 11th IEEE International Conference on*, pp. 573–575. <http://dx.doi.org/10.1109/ICALT.2011.173>.
- [71] Saaty, T.L., 1980. *The Analytic Hierarchy Process*. McGraw-Hill, N. York, USA.
- [72] Sachin, R., Vijay, M., 2012. A survey and future vision of data mining in educational field. In: *Advanced Computing Communication Technologies (ACCT), 2012, Second International Conference on*, pp. 96–100. <http://dx.doi.org/10.1109/ACCT.2012.14>.
- [73] Sankar, C.S., Clayton, H., 2010. An evaluation of use of multimedia case studies to improve an introduction to information technology course. *Int. J. Inf. Commun. Technol. Educ.* 6 (3), 25–37. <http://dx.doi.org/10.4018/jicte.2010070103>. ISSN 1550-1876.
- [74] Santos, O., Boticario, J., 2008. Improving learners' satisfaction in specification-based scenarios with dynamic inclusive support. In: *Advanced Learning Technologies, 2008. ICALT '08. Eighth IEEE International Conference on*, pp. 491–495. <http://dx.doi.org/10.1109/ICALT.2008.101>.
- [75] Scheffer, T., 2001. Finding association rules that trade support optimally against confidence. In: *Proceedings of the 5th European Conference on Principles of Data Mining and Knowledge Discovery, PKDD '01*. Springer-Verlag, London, UK, pp. 424–435 ISBN 3-540-42534-9.
- [76] Shahiri, A.M., Husain, W., Rashid, N.A., 2015. A review on predicting student's performance using data mining techniques. *Proc. Comput. Sci.* 72, 414–422. <http://dx.doi.org/10.1016/j.procs.2015.12.157>. ISSN 1877-0509. The Third Information Systems International Conference.
- [77] Shen, R., Tang, Y., Zhang, T., 2001. The intelligent assessment system in web-based distance learning education. In: *Frontiers in Education Conference, 2001. 31st Annual*. vol. 1. pp. 7–11. <http://dx.doi.org/10.1109/FIE.2001.963855>.
- [78] Song, D., Lin, H., Yang, Z., 2007. Opinion mining in e-learning system. In: *Network and Parallel Computing Workshops, 2007. NPC Workshops. IFIP International Conference on*, pp. 788–792. <http://dx.doi.org/10.1109/NPC.2007.51>.

- [79] Sun, X., Zhao, W., 2009. Design and implementation of an e-learning model based on wum techniques. In: E-Learning, E-Business, Enterprise Information Systems, and E-Government, 2009. EEEE '09. International Conference on, pp. 248–251. <http://dx.doi.org/10.1109/EEEE.2009.88>.
- [80] Tang, N., Chen, Z., Lin, T., Cui, J., 2012. Research and design of applying association rule in course management system. In: Consumer Electronics, Communications and Networks (CECNet), 2012 2nd International Conference on, pp. 3394–3398. <http://dx.doi.org/10.1109/CECNet.2012.6201470>.
- [81] Tovar, E., Soto, O., 2010. The use of competences assessment to predict the performance of first year students. In: Frontiers in Education Conference (FIE), 2010 IEEE, <http://dx.doi.org/10.1109/FIE.2010.5673566>. pp. F3J-1-F3J-4.
- [82] Wang, C.-S., Lin, S.-L., 2012. Combining fuzzy ahp and association rule to evaluate the activity processes of e-learning system. In: Genetic and Evolutionary Computing (ICGEC), 2012 Sixth International Conference on, pp. 566–570. <http://dx.doi.org/10.1109/ICGEC.2012.72>.
- [83] Wang, J., Lu, Z., Wu, W., Li, Y., 2012. The application of data mining technology based on teaching information. In: Computer Science Education (ICCSE), 2012 7th International Conference on, pp. 652–657. <http://dx.doi.org/10.1109/ICCSE.2012.6295159>.
- [84] Wang, L., Li, J., Ding, L., Li, P., 2010. E-learning evaluation system based on data mining. In: Information Engineering and Electronic Commerce (IEEC), 2010 2nd International Symposium on, pp. 1–3. <http://dx.doi.org/10.1109/IEEC.2010.5533204>.
- [85] Wang, L., Meinel, C., 2007. Mining the students' learning interest in browsing web-streaming lectures. In: Computational Intelligence and Data Mining, 2007. CIDM 2007. IEEE Symposium on, pp. 194–201. <http://dx.doi.org/10.1109/CIDM.2007.368872>.
- [86] Wang, W., Maruatona, O., Qian, H., 2011. Video games' educational evaluation model based on bp neural network. In: Complexity and Data Mining (IWCDM), 2011 First International Workshop on, pp. 156–158. <http://dx.doi.org/10.1109/IWCDM.2011.44>.
- [87] Wang, X., Yang, D., Wen, M., Koedinger, K., Rose, C., 2015. Investigating how student's cognitive behavior in mooc discussion forum affect learning gains. In: Santos, O.C., Boticario, J.G., Romero, C., Pechenizkiy, M., Merceron, A., Mitros, P., Luna, J.M., Mihaescu, C., Moreno, P., Hershkovitz, A., Ventura, S., Desmarais, M. (Eds.), Proceedings of the 8th International Conference on Educational Data Mining. International Conference on Educational Data Mining (EDM), International Educational Data Mining Society (IEDMS), Madrid, Spain, pp. 226–233 ISBN 978-84-606-9425-0.
- [88] Xing, W., Guo, R., Petakovic, E., Goggins, S., 2015. Participation-based student final performance prediction model through interpretable genetic programming: Integrating learning analytics, educational data mining and theory. *Comput. Human Behav.* 47, 168–181.

- <http://dx.doi.org/10.1016/j.chb.2014.09.034>. ISSN 0747–5632. Learning Analytics, Educational Data Mining and data-driven Educational Decision Making.
- [89] Yu, J., Sun, B., 2010. Research of education evaluation information mining technology based on analytical hierarchy process (ahp) and genetic algorithm (ga). In: Advanced Computer Theory and Engineering (ICACTE), 2010 3rd International Conference on. vol. 4 <http://dx.doi.org/10.1109/ICACTE.2010.5579725>. pages V4-289-V4-292.
- [90] Yu, T., Jo, I.-H., 2014. Educational technology approach toward learning analytics: Relationship between student online behavior and learning performance in higher education. In: Proceedings of the Fourth International Conference on Learning Analytics And Knowledge, LAK '14. ACM, New York, NY, USA, pp. 269–270. <http://dx.doi.org/10.1145/2567574.2567594>. ISBN 978-1-4503-2664-3.
- [91] Yubing, A., Jianping, Z., 2010. The application of data mining technology in distance learning evaluation. In: Information Technology and Applications (IFITA), 2010 International Forum on. vol. 3. pp. 145–148. <http://dx.doi.org/10.1109/IFITA.2010.153>.
- [92] Zaiane, O., Luo, J., 2001. Towards evaluating learners' behaviour in a web-based distance learning environment. In: Advanced Learning Technologies, 2001. Proceedings. IEEE International Conference on, pp. 357–360. <http://dx.doi.org/10.1109/ICALT.2001.943944>.
- [93] Zeng, L.-Y., 2012. An evaluation system of game-based learning based on data mining. In: Computer Science and Network Technology (ICCSNT), 2012 2nd International Conference on, pp. 1732–1736. <http://dx.doi.org/10.1109/ICCSNT.2012.6526255>.
- [94] Recommender systems the textbook (2016) – Charu C. Aggarwal
- [95] The Definition of Novelty in Recommendation System - Liang Zhang
- [96] Practical Recommender systems – Kim Falk