

**ВЕРОЯТНОСТНЫЕ ЗАКОНОМЕРНОСТИ В ЦИФРОВЫХ  
СЛЕДАХ ПОЛЬЗОВАТЕЛЕЙ СИСТЕМЫ ЭЛЕКТРОННОГО  
ОБУЧЕНИЯ УЧРЕЖДЕНИЯ ОБРАЗОВАНИЯ «БЕЛОРУССКИЙ  
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИНФОРМАТИКИ  
И РАДИОЭЛЕКТРОНИКИ»**

*Крез К.С., Дроздова М.А.*

*Белорусский государственный университет информатики и радиоэлектроники,  
г. Минск, Республика Беларусь*

*Научный руководитель: Шнейдеров Е.Н. – канд.техн.наук, доцент, доцент кафедры ПИКС,  
проректор по учебной работе*

**Аннотация.** Одной из задач в области интеллектуального анализа образовательных данных (*EDM*) является поиск закономерностей, позволяющих количественно оценить эффективность обучения с использованием систем электронного обучения (*LMS*). В статье рассмотрено влияние работы пользователей системы электронного обучения БГУИР с теоретическим материалом по учебным дисциплинам на получаемые ими отметки по практическим работам, а также алгоритм получения статистических данных для оценки степени этого влияния. Результатом является формулировка общих выводов, позволяющих определить дальнейшие направления деятельности в исследуемой области.

**Ключевые слова:** система электронного обучения, модель, цифровой след, анализ данных, закономерности

В процессе функционирования системы электронного обучения учреждения образования (далее – СЭО) на базе *Moodle LMS* каждый пользователь оставляет в ней так называемый «цифровой след». Цифровой след представляет собой, как правило, информацию о самом пользователе, различные записи в базах данных (о деятельности, о сообщениях, об оценках и др.), отдельные файлы и другие данные. В неявном виде цифровой след пользователей также включает информацию о методике его работы с учебными материалами, а значит, методике его обучения применительно к конкретной дисциплине [1].

Основной целью исследования в рамках статьи является проверка гипотезы (утверждения) «чем чаще студенты дистанционной формы получения образования просматривают теоретические материалы по определённым темам до выполнения практического задания по ним, тем выше получаемая студентами отметка по этому заданию» в цифровых следах пользователей системы электронного обучения учреждения образования «Белорусский государственный университет информатики и радиоэлектроники» (далее – БГУИР), организованной на базе *Moodle LMS*.

Подтверждение гипотезы в указанном случае основано на наличии закономерностей в виде значимой линейной корреляции между количеством просмотров теоретических материалов до выполнения практического задания и отметкой за практическое задание (далее – коэффициенты корреляции). Значимость корреляции может быть определена с использованием шкалы Чеддока.

В качестве примера рассмотрим пошаговое получение значений коэффициентов корреляции для дисциплины «Метрология, стандартизация и сертификация в радиоэлектронике»:

1. С использованием *SQL*-запроса получаем информацию о деятельности студентов БГУИР дистанционной формы обучения с сентября 2021 года по январь 2022 года применительно ко всем изучаемым ими учебным дисциплинам. Эта информация, представляющая собой поведенческие данные, была выгружена в промежуточный файл формата \*.csv. В указанном формате она занимает около 2 Гб.

2. Далее \*.csv файл был импортирован в *JupyterLab* для анализа. Основная рассматриваемая для получения данных таблица – *logstore\_standart log*, хранящая, так называемый, «журнал событий», и таблица *grade\_grades*, хранящая «отметки». Таблица «журнал событий» содержит следующие основные поля: *id*, *eventname*, *component*, *action*, *target*, *userid*, *courseid*, *timecreated* (рисунок 1), таблица «отметки» содержит такие поля: *id*, *userid*, *finalgrade* (рисунок 2).

	id	eventname	component	action	target	objecttable	objectid	crud	edulevel	contextid	...
0	66313229	\core\event\user_loggedin	core	loggedin	user	user	9254.0	r	0	1	...
1	66313230	\core\event\course_viewed	core	viewed	course	NaN	NaN	r	2	2	...
2	66313231	\core\event\course_viewed	core	viewed	course	NaN	NaN	r	2	3371	...
3	66313232	\mod_page\event\course_module_viewed	mod_page	viewed	course_module	page	10374.0	r	2	271545	...
4	66313233	\core\event\course_viewed	core	viewed	course	NaN	NaN	r	2	1349	...
...	...	...	...	...	...	...	...	...	...	...	...
1716974	94967232	\core\event\notification_viewed	core	viewed	notification	notifications	1510864.0	u	0	42699	...
1716975	94967233	\core\event\course_viewed	core	viewed	course	NaN	NaN	r	2	246860	...
1716976	94967234	\core\event\webservice_function_called	core	called	webservice_function	NaN	NaN	r	0	1	...
1716977	94967235	\core\event\webservice_function_called	core	called	webservice_function	NaN	NaN	r	0	1	...
1716978	94967236	\core\event\webservice_function_called	core	called	webservice_function	NaN	NaN	r	0	1	...

Рисунок 1 – Фрагмент таблицы «журнал событий».

	id	timecreated	items	userid	courseid	finalgrade
2098	2022-05-31 12:25:33	quiz	2099	296	9	
2099	2022-05-31 13:09:59	quiz	2099	296	8	
2100	2022-05-31 13:19:24	quiz	2099	296	10	
2101	2022-05-31 13:22:01	quiz	2099	296	8	
3564	2022-06-05 22:06:09	assign	2099	296	0	
2098	2022-06-04 12:49:38	quiz	2126	296	8	
2099	2022-06-04 13:21:39	quiz	2126	296	8	
2101	2022-06-04 13:30:28	quiz	2126	296	8	
3564	2022-06-04 15:43:08	assign	2126	296	0	
2098	2022-06-02 13:37:40	quiz	10554	296	6	

Рисунок 2 – Фрагмент таблицы «отметки»

Задачей анализа указанных таблиц является получение событий о взаимодействии пользователей с элементами учебной дисциплины. Такие события возникают, если обучающийся просмотрел страницу, скачал файл, написал на форуме, отправил сообщение, получил оценку и т. д. Для уточнения (фильтрации) данных необходимы также и другие таблицы (*role* и *role\_assignments* – для выборки студентов исключительно дистанционной формы, *course* – для выборки студентов исключительно дистанционной формы).

Для разделения материала дисциплины на теоретический и практический используется поле *objecttable* таблицы «журнал событий». Теоретический материал представляет собой элементы *data*, *bigbluebutton*, *wiki*, *url*, *glossary*, *book*, *lesson*, *folder*, *label*, *page*, *chat* и *resource*. Практический материал – *assign* (задание), *quiz* (тест) и *vpl* (виртуальная лаборатория программирования).

3. С использованием библиотек *numpy* и *pandas* для удобства полученная информация была сведена в единую таблицу (рисунок 3). Она содержит время просмотра отдельными студентами дистанционной формы обучения теоретических материалов, время выполнения ими практических заданий, а также полученные за практические задания отметки.

	timecreated	items	userid	courseid	finalgrade
2	2021-09-01 00:00:04	viewed	1909	126	NaN
3	2021-09-01 00:00:10	viewed	1909	126	NaN
4	2021-09-01 00:00:11	viewed	7876	59	NaN
9	2021-09-01 00:00:15	viewed	7876	59	NaN
6	2021-09-01 00:00:15	viewed	7876	59	NaN
...	...	...	...	...	...
1716708	2022-06-30 23:28:44	viewed	14280	271	NaN
1716709	2022-06-30 23:28:48	viewed	14280	271	NaN
1716710	2022-06-30 23:28:52	viewed	14280	271	NaN
1716832	2022-06-30 23:47:07	viewed	15879	4882	NaN
1716836	2022-06-30 23:47:13	viewed	15879	4882	NaN

Рисунок 3 – Сводная таблица

4. Для получения значений коэффициентов линейной корреляции применительно к отдельным дисциплинам будем использовать метод *np.corrcoef()* библиотеки *numpy*, который использует классическую формулу 1:

$$r = \frac{\sum_{i=1}^n (X_i - \underline{X})(Y_i - \underline{Y})}{\sqrt{\sum_{i=1}^n (X_i - \underline{X})^2 \sum_{i=1}^n (Y_i - \underline{Y})^2}} \quad (1)$$

где  $X$  – количество просмотров отдельными студентами теоретического материала в СЭО до выполнения ими практической работы;

$Y$  – отметка, полученная в СЭО этими студентами за выполнение практической работы;

$n$  – количество отметок, полученных студентами в отдельной рассматриваемой дисциплине.

Для удобства вычисления значений по формуле (1) данные сводной таблицы (рисунок 3), были трансформированы в таблицу «просмотры-отметка», фрагмент которой изображён на рисунке 4.

	Mark	Theory		Mark	Theory
0	9.0	36.0	0	9.0	36.0
1	8.0	29.0	1	8.0	65.0
2	8.0	34.0	2	8.0	99.0
3	9.0	25.0	3	9.0	124.0
4	9.0	21.0	4	9.0	145.0
5	10.0	35.0	5	10.0	180.0
6	9.0	117.0	6	9.0	297.0
7	9.0	23.0	7	9.0	320.0
8	0.0	74.0	8	0.0	394.0

Рисунок 4 – Фрагмент таблицы «просмотры-отметка»

В таблице «просмотры-отметка» столбец *theory* содержит значения множества  $X$ , а столбец *mark* – значения множества  $Y$ .

5. Полученные на 4 шаге значения коэффициентов линейной корреляции для каждого студента дисциплины «Метрология, стандартизация и сертификация в радиоэлектронике» сведены в таблицу «коэффициенты», фрагмент которой изображён на рисунке 5.

	userid	correlation
0	2577.0	NaN
1	22856.0	NaN
2	8904.0	NaN
3	2724.0	NaN
4	2126.0	0.710542
5	2099.0	0.708593
6	8908.0	NaN
7	2114.0	NaN
8	17351.0	0.855072
9	10554.0	0.178971
10	2735.0	NaN

Рисунок 5 – Фрагмент таблицы «коэффициенты»

В таблице «коэффициенты» столбец *userid* содержит идентификатор пользователя, применительно к которому получено значение, а также значение коэффициента линейной корреляции для него. Можно заметить, что в таблице присутствуют *NaN*-значения, которые могут быть объяснены недостаточностью данных для вычисления значения по формуле (1). При дальнейших вычислениях *NaN*-значения не учитываются (исключаются из таблицы).

6. Для вычисления значения коэффициента линейной корреляции применительно ко всей дисциплине «Метрология, стандартизация и сертификация в радиоэлектронике» используется формула 2:

$$\bar{r} = \frac{1}{N} \sum_{i=1}^N r_i \quad (2)$$

где  $r_i$  – значение вычисленного коэффициента линейной корреляции для  $i \in [1, N]$ ;  $N$  – количество полученных значений таблицы «коэффициенты».

Для рассматриваемой дисциплины значение коэффициента линейной корреляции по формуле (2) составило  $r = 0,6133$ , что по шкале Чеддока классифицируется как средняя степень связи. Это говорит о наличии вероятностной зависимости между количеством просмотров студентами теоретических материалов по определённым темам дисциплины до выполнения практического задания и отметкой, полученной ими по этому заданию. Применительно к дисциплине «Метрология, стандартизация и сертификация в радиоэлектронике» гипотеза подтверждается.

Аналогично по шагам 1...5 были вычислены коэффициенты корреляции для 35 учебных дисциплин, изучаемых студентами дистанционной формы получения образования. Некоторые полученные значения приведены в таблице 1.

Таблица 1 — Значения коэффициентов линейной корреляции для некоторых учебных дисциплин студентов дистанционной формы получения образования

Название учебной дисциплины	Коэффициент линейной корреляции
1. Отраслевой маркетинг	0,9878
2. Организация информационно-компьютерных систем и сетей. Часть 2	0,9723
3. Избранные главы информатики	0,9313
...	...
34. Электронные приборы	0,0385
35. Программное обеспечение мобильных систем. Часть 1	0,0026

Интервал полученных значений коэффициентов линейной корреляции от 0,0026 до 0,9878. Среднее значение коэффициента линейной корреляции составляет 0,45.

В результате исследования можно сделать следующие выводы:

– использование коэффициента линейной корреляции не даёт представления о деталях образовательного процесса, мотивации студентов, однако позволяет количественно оценить влияние теоретических материалов на получаемые студентами дистанционной формы получения образования отметки;

– широкий разброс коэффициентов корреляции, а также наличие *NaN*-значений в таблице «коэффициенты» может говорить о том, что ключевым фактором, влияющим на значения коэффициентов, является качество представления и содержания теоретических материалов в учебных дисциплинах;

– учитывая множественность шагов обработки данных (видно из примера расчётов для дисциплины «Метрология, стандартизация и сертификация в радиоэлектронике») для выполнения анализа цифровых следов эффективно использовать отдельную реляционную базу данных, содержащую только необходимую изолированную информацию о цифровых следах пользователей и оптимизированную для операций выборки, фильтрации и объединения.

### Список литературы.

1. Аминов Т.К. Цифровой след, как средство развития образовательной деятельности и модернизации учебных программ / Т.К. Аминов, А.С. Волков, Е.В. Желнина // Актуальные проблемы гуманитарных и социально-экономических наук. – 2022. – № 2(85). – с. 7–14. UDC 621.3.049.77–048.24:537.2

## CONTROL OF MICROCONTROLLER UNDER THE INFLUENCE OF ELECTROSTATIC DISCHARGE

*Krez K.S., Drozdova M.A.*

*Belarusian State University of Informatics and Radioelectronics, Minsk, Republic of Belarus*

*Shneiderov E.N. – PhD, associate professor, associate professor of the Department of ICSD, vice-rector for academic work*

**Annotation.** One of the tasks in the field of intellectual analysis of educational data (EDM) is the search for patterns that allow quantifying the effectiveness of learning using e-learning systems (LMS). The article examines the impact of the work of users of the BSUIR e-learning system with theoretical material on academic disciplines on the marks they receive on practical work, as well as an algorithm for obtaining statistical data to assess the degree of this influence. The result is the formulation of general conclusions that allow us to determine further areas of activity in the field under study.

**Keywords:** e-learning system, model, digital footprint, data analysis, patterns