**Аналитический отчёт**

**Тематическое моделирование**

**Жаппарова Амина**

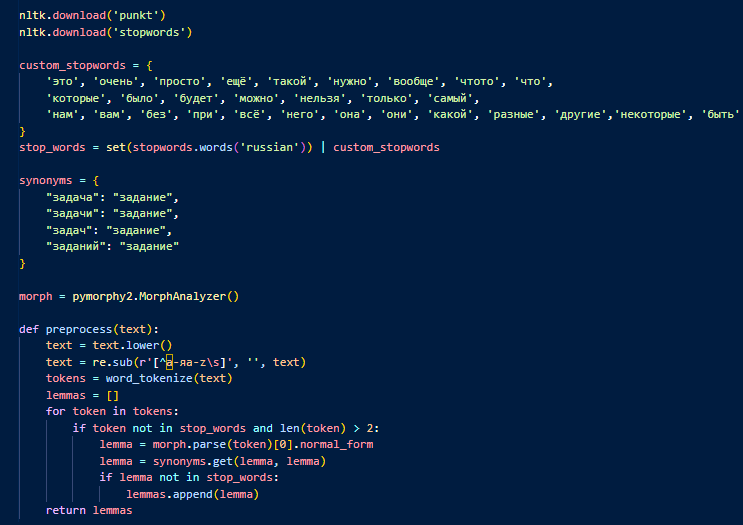
**Общая цель:**

Понять основные темы и проблемные/сильные стороны онлайн-курсов на основе текстов позитивных и негативных отзывов с применением двух подходов:

* LDA (Latent Dirichlet Allocation)
* BERTopic

**Предобработка текста:**

Приведение к нижнему регистру, удаление лишних символов, токенизация и лемматизация (pymorphy2), удаление стандартных и пользовательских стоп-слов (nltk), замена синонимов (например, "задача" = "задание").



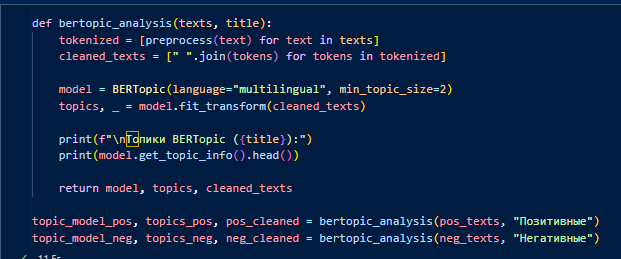
**Модели:**

* LDA с числом тем = 5 для позитивных и негативных отзывов

Изображение выглядит как текст, снимок экрана, Шрифт, программное обеспечение

Содержимое, созданное искусственным интеллектом, может быть неверным.

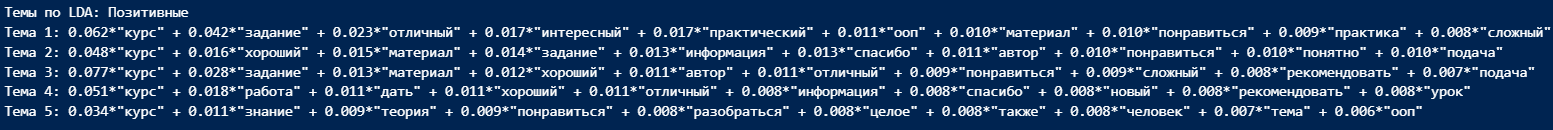
* BERTopic с автоматическим определением числа тем (минимум 2)



**Результаты:**

**Позитивные отзывы**

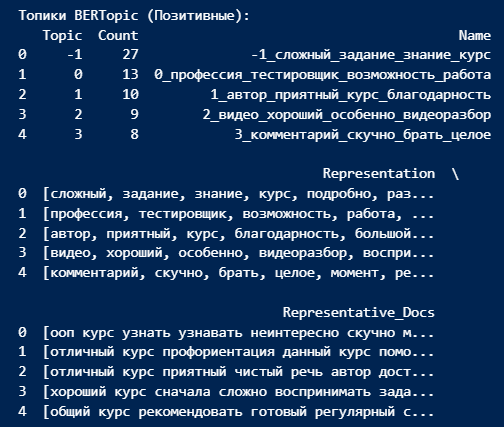
**LDA Темы и ключевые слова:**

****

1. Курс, задание, отличный, интересный, практический
2. Курс, материал, спасибо, автор, понятно
3. Курс, хороший, автор, рекомендовать, подача
4. Курс, работа, информация, новый, урок
5. Знание, теория, разобраться, тема

Интерпретация:   
Курсы вызывают интерес, практичны, содержат хорошие задания. Хорошо структурированные, полезные материалы, благодарность автору. Подача материала важна, автор — значимая фигура. Практическое применение знаний и полезность. Курсы помогают разобраться в теории, расширяют кругозор.

**BERTopic Темы:**

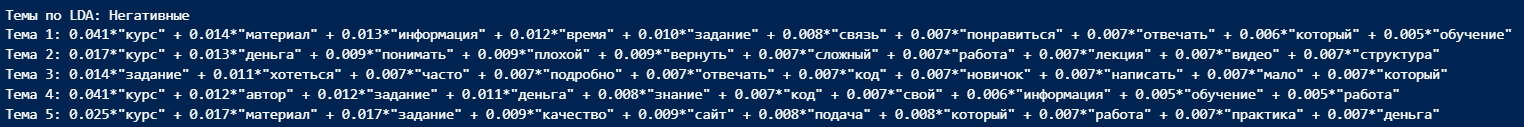
****

1. сложный\_задание\_знание\_курс
2. профессия\_тестировщик\_работа
3. автор\_приятный\_благодарность
4. видео\_видеоразбор
5. комментарий\_целое

Интерпретация:  
Курс сложный, но обучающий. Польза курса в профессиональном развитии. Уважение к преподавателю и благодарность. Высокая ценность видеоматериалов. Полезные замечания, целостность курса.

**Негативные отзывы**

**LDA Темы:**

****

1. Материал, информация, время, связь
2. Деньга, плохой, вернуть, видео
3. Задание, новичок, код, мало
4. Автор, задание, знание, обучение
5. Материал, подача, сайт

Интерпретация:  
Недостаток содержательности, плохая структурность. Жалобы на соотношение цены-качества. Неподходящий уровень, особенно для новичков. Нарекания к автору, заданиям и общему обучающему эффекту. Критика интерфейса, подачи и платформы в целом.

**BERTopic Темы:**

**Изображение выглядит как текст, снимок экрана, Шрифт, меню

Содержимое, созданное искусственным интеллектом, может быть неверным.**

1. курс\_задание\_автор\_дизайн
2. задание\_домашний\_писать\_странный
3. навык\_желающий\_разработчик\_курс
4. искать\_знание\_приходиться\_сложность
5. место\_понравиться\_корочка\_решить

Общая критика курса: дизайн, автор, задания. Задания странные и непродуманные. Курс не подходит для получения реальных навыков. Приходится искать информацию вне курса, чрезмерная сложность. Мотивация: “получить сертификат”, а не знания

**Рекомендации:**

1. Пересмотреть цены на курсы
2. Обеспечить стабильную обратную связь от преподавателей, кураторов
3. Сделать курсы более ориентированными на реальные задачи
4. Перепроверить задания на логичность и адаптировать под уровень участников
5. Сделать структуру курса более связной и доступной