Федеральное государственное автономное образовательное учреждение высшего образования
"Национальный исследовательский университет
"Высшая школа экономики"

Факультет коммуникаций, медиа и дизайна
Департамент интегрированных коммуникаций

Рабочая программа дисциплины
Проектно-исследовательский семинар
«Анализ данных в коммуникационных проектах»

для образовательной программы «Коммуникации, основанные на данных»
направления подготовки 42.04.01 Реклама и связи с общественностью
уровень магистр

Разработчики программы:
Грызунова Елена Аркадьевна, кандидат социологических наук, egryzunova@hse.ru
Егоров Дмитрий Владимирович, d.egorov@weborama.com.ru
Суворина Татьяна Владимировна, t.suvorina@weborama.com.ru

Академический руководитель образовательной программы
Е. А. Грызунова

Москва, 2018

Настоящая программа не может быть использована другими подразделениями университета и другими вузами без разрешения подразделения-разработчика программы.

1. Область применения и нормативные ссылки

Настоящая программа учебной дисциплины устанавливает требования к образовательным результатам и результатам обучения студента и определяет содержание и виды учебных занятий и отчетности.

Программа предназначена для преподавателей, ведущих дисциплину «Проектно-исследовательский семинар «Анализ данных в коммуникационных проектах», учебных ассистентов и студентов направления подготовки 42.04.01 Реклама и связи с общественностью, обучающихся по образовательной программе «Коммуникации, основанные на данных».

Программа учебной дисциплины разработана в соответствии с:

- Образовательным стандартом НИУ ВШЭ по направлению 42.04.01 "Реклама и связи с общественностью" подготовки магистра (ред. 2017 г.);
- Образовательной программой «Коммуникации, основанные на данных» направления 42.04.01 "Реклама и связи с общественностью" подготовки магистра.
- Объединенным учебным планом университета по образовательной программе «Коммуникации, основанные на данных», утвержденным в 2018 г.

2. Цели освоения дисциплины

Целью освоения дисциплины «Проектно-исследовательский семинар «Анализ данных в коммуникационных проектах» является освоение студентами аналитических и управленческих компетенций, необходимых для реализации прикладных коммуникационных проектов, основанных на данных.

3. Компетенции обучающегося, формируемые в результате освоения дисциплины

Уровни формирования компетенций:

РБ — ресурсная база, в основном теоретические и предметные основы (знания, умения);

СД – способы деятельности, составляющие практическое ядро данной компетенции;

МЦ – мотивационно-ценностная составляющая, отражает степень осознания ценности компетенции человеком и готовность ее использовать

В результате освоения дисциплины студент осваивает компетенции:

Компетенция	Код по ОС НИУ ВШЭ	Уровень формирования компетенции	Дескрипторы – основные признаки освоения (показатели достижения результата)	Формы и методы обучения, способствующие формированию и развитию компетенции	Форма контроля уровня сформированности компетенции
Способен рефлексировать (оценивать и перерабатывать) освоенные научные методы и способы деятельности.	УК-1	РБ, СД	Дает определение методам, которые использует в работе с данными, понимает принцип работы и способы модификации для достижения определенных результатов и предлагать новые решения	Семинары, выполнение домашних заданий	Выполнение практических домашних заданий по анализу данных. Разбор кейсов. Защита плана курсового проекта, предзащита группового проекта и курсового проекта.
Способен создавать новые теории, изобретать новые способы и инструменты профессиональной деятельности.	УК-2	РБ, СД	Понимает стандартные алгоритмы и методы анализа данных для решения прикладных задач, способен их комбинировать и разрабатывать новые подходы	Семинары, выполнение домашних заданий	Выполнение практических домашних заданий по анализу данных. Подготовка к выполнению мини-проекта. Защита плана курсового проекта, предзащита группового проекта и курсового проекта.
Способен совершенствовать и развивать свой интеллектуальный и культурный уровень, строить траекторию профессионального развития и карьеры.	УК-4	РБ, СД, МЦ	Принимает участие в дискуссиях с представителями индустрии, разбирает кейсы, предлагает идеи для решения прикладных задач и выполнения мини-проекта	Семинары, домашние задания	Разбор кейсов, плана курсового проекта, предзащита группового проекта и курсового проекта.
Способен вести профессио-	УК-8	РБ, СД	Способнен оценивать результаты	Семинары, выполнение домашних	Экзамен, оригинальность ис-

Компетенция	Код по ОС НИУ ВШЭ	Уровень формирования компетенции	Дескрипторы – основные признаки освоения (показатели достижения результата)	Формы и методы обучения, способствующие формированию и развитию компетенции	Форма контроля уровня сформированности компетенции
нальную, в том числе научно-исследовательскую деятельность в международной среде.			исследований, международных стандартов по big data и анализу данных. Исследует зарубежные практики анализа данных в коммуникационных проектах	заданий,	следования в рамках курсовой работы и выполнения других проектов по ходу курса.
Способен разработать и методически использовать концептуальные модели, рабочие планы и программы проведения научных и проектных исследований в области рекламы и связей с общественностью.	ОПК-1	РБ, СД	Понимает основные направления современных научных и проектных исследований, знает, какие данные нужны, как их собрать и с помощью чего обработать для проверки гипотезы, предлагает собственные гипотезы и составляет план проекта или научного исследования для решения своих задач	Семинары, практические задачи, выполнение домашних заданий	Презентация плана курсового проекта, выполнение домашних заданий и мини-проекта.
Способен применять, творчески развивать накопленные знания и опираться на них в исследовательской деятельности.	ОПК-2	РБ, СД, МЦ	Владеет основными методами анализа данных, понимает, как выбрать нужный и оптимизировать под конкретную исследовательскую задачу. Находит необходимую научную информацию в Сети.	Семинары, практические задачи	Выполнение практических домашних заданий по анализу данных. Разбор кейсов.

Компетенция	Код по ОС НИУ ВШЭ	Уровень формирования компетенции	Дескрипторы – основные признаки освоения (показатели достижения результата)	Формы и методы обучения, способствующие формированию и развитию компетенции	Форма контроля уровня сформированности компетенции
Способен проводить сравнительный анализ возможностей различных медиа-каналов для выхода на аудиторию с использованием наиболее подходящих медиа-носителей	ПК-2	РБ, СД, МЦ	Понимает и распознает индикаторы эффективной коммуникации, владеет инструментами измерения и увеличения эффективности, способен проводить медиа и аудиторное планирование, провести анализ эффективности текуших каналов и выбрать наиболее подходящие	Семинары, практические занятия по оценке эффективности рекламных каналов	Экзамен, домашние задания
Способен составлять аналитические справки, экспертные заключения, отчеты, обзоры и прогнозы на основе результатов анализа коммуникационной среды	ПК-3	РБ, СД, МЦ	Владеет и применяет методы анализа данных для анализа аудитории, составления портрета пользователя, оценки эффективности кампаний. Способен представить результате в формате презентации, отчета или интерактивного дэшборда для презентации клиенту.	Семинары, практические занятия, выполнение домашних заданий	Презентация проектов студентов.
Способен обрабатывать данные, с целью построения коммуникационной кампании, в том числе используя специальное программное обеспечение	ПК-5	РБ, СД, МЦ	Демонстрирует владение статистическим анализом данных с использованием Python, демонстрирует владение инструментами визуализации данных с использованием Python/BI инструментов/презентаций	Семинары, практические занятия, выполнение домашних заданий	Выполнение домашних заданий и демонстрация проектов.

Компетенция	Код по ОС НИУ ВШЭ	Уровень формирования компетенции	Дескрипторы – основные признаки освоения (показатели достижения результата)	Формы и методы обучения, способствующие формированию и развитию компетенции	Форма контроля уровня сформированности компетенции
Способен к презентации и популяризации исследовательских и инструментальных навыков в области рекламы и связей с общественностью	ПК-8	РБ, СД	Способен объяснить научно-популярным языком принципы работы алгоритма, цели и задачи проекта, донести до клиентов, выступать на конференциях с кейсами	Семинары, практические занятия, выполнение домашних заданий	Демонстрация проектов перед однокурсниками, разбор кейсов
Способен ставить задачи исследования, выбирать методы экспериментальной работы, подготавливать базу для научных исследований	ПК-10	РБ, СД, МЦ	Понимает принципы работы алгоритмов анализа данных, знает отличия от стандартных статистических подходов, понимает, какую модель выбрать для решения конкретной задачи и какие данные для этого необходимо собрать и каким образом собрать их безопасным способом	Семинары, практические занятия, выполнение домашних заданий	Выполнение домашних заданий и демонстрация проектов.

4. Место дисциплины в структуре образовательной программы

Настоящая дисциплина относится к блоку дисциплин «Практика(и), проектная и(или) научно-исследовательская работа». Для программы «Коммуникации, основанные на данных» данная дисциплина является обязательной.

Изучение данной дисциплины базируется на следующих дисциплинах:

- Основы дата-ориентированных коммуникаций
- Программирование для анализа данных

Для освоения учебной дисциплины студенты должны владеть следующими знаниями и компетенциями:

Понимание экосистемы коммуникаций, основанных на данных

Базовые навыки в области программирования, знание основных методов анализа данных

Основные положения дисциплины должны быть использованы в дальнейшем при изучении дисциплин:

- Проектно-исследовательский семинар «Управление проектами по методологии CRISP-DM»
- Большие данные в коммуникационных стратегиях

5. Тематический план учебной дисциплины

Тематический план 2 модуля

№	Название раздела	Всего часов	Аудиторные часы				Самостоятельная работа
			Лекции	Семинары	Практические занятия	Другие виды работы	
1	Экосистема рекламного (digital) рынка и его эволюция	12		4			8
2	Презентация и обсуждение предложенных руководителями тем курсовых проектов	12		4			8
3	Индикаторы эффективной коммуникации	12		4			8
4	Инструменты измерения эффективности	12		4			8
5	Инструменты увеличения эффективности	12		4			8
6	Онлайн данные, оффлайн данные, онлайн+ оффлайн	12		4			8
7	Смежные источники данных и аналитики	12		4			8
8	От медиа планирования к аудиторному планированию	12		4			8
9	Безопасность данных	12		4			8

10	От обезличенной коммуникации к персональной	12		4			8
11	Кейсы из индустрии (приглашенные спикеры)	12		4			8
12	Анализ данных для решения задач рекламодателей	12		4			8
13	Первичный анализ данных	10		2			8
14	Задачи обучения с учителем. Как построить свой Look-a-like.	16		4			12
15	Обучение без учителя. Поиск трендов и групп среди аудитории сайта	16		4			12
16	Другие прикладные задачи анализа данных	12		4			8
17	Оценка эффективности рекламных кампаний	12		4			8
18	Визуализация результатов	12		4			8
19	Демонстрация мини-проектов	16		4			12
	Итого по 2 модулю	**238**		**74**			**164**

Тематический план 3 модуля

№	Название раздела	Всего часов	Аудиторные часы				Самостоятельная работа
			Лекции	Семинары	Практические занятия	Другие виды работы	
1	Введение в проектную работу	**14**		**4**			**10**
2	Построение и проверка гипотез в интеллектуальном анализе данных	**30**		**10**			**20**
3	Презентация и обсуждение планов курсовых работ	**40**		**10**			**30**
4	Предзащита и обсуж-	**36**		**6**			**30**

	дение групповых проектов						
	Итого по 3 модулю	**120**		**30**			**90**

Тематический план 4 модуля

№	Название раздела	Всего часов	Аудиторные часы				Самостоятельная работа
			Лекции	Семинары	Практические занятия	Другие виды работы	
1	Интеллектуальный анализ текста	**48**		**20**			**28**
2	Введение в сетевой анализ	**30**		**12**			**18**
3	Предзащита курсовых работ	**20**		**8**			**12**
	Итого по 4 модулю	**98**		**40**			**58**

6. Формы контроля знаний студентов

Тип кон-кон-троля	Форма контроля	1 год			Параметры **
		2	3	4	
Текущий	Контрольная работа				
	Эссе				
	Реферат				
	Коллоквиум				
	Домашнее задание	*			3 практических задания по теме семинаров
	Самостоятельная работа				
	Лабораторная работа				
	Проект	*			Мини-проект по 2 модулю (можно выбрать расширенную версию одного из домашних заданий или предложить свою идею) и его защита

	Презентация плана курсового проекта		*		Презентация от 10 слайдов на 10 минут, ответы на вопросы преподавателя и других студентов по курсовому проекту.
	Предзащита группового проекта		*		Презентация от 15 слайдов на 15 минут, ответы на вопросы преподавателя и других студентов по групповому проекту.
	Предзащита курсовых проектов			*	Презентация от 10 слайдов на 10 минут, ответы на вопросы преподавателя по курсовому проекту.
Итоговый	Экзамен			*	Устный экзамен по билетам, 2 вопроса в билете.

7. Критерии оценки знаний, навыков

Оценки по всем формам текущего контроля выставляются по 10-ти балльной шкале.

2 модуль

Домашние задания.

Учащимся предлагается три практических задания по анализу аудитории и составлению статистических сравнений характеристик пользователей, построению базовых скоринговых моделей и кластеризации пользователей в виде тетрадок Jupyter, в которых требуется дописать свой код, и примерный набор данных к ним. В результате студенты получают заготовку для будущих исследований и лучше разбираются в алгоритмах подготовки данных, обучения и валидации моделей, формирования конечного результата и оценки качества алгоритма, учатся визуализировать результаты.

Результатом домашней работы должна быть заполненная тетрадка Jupiter Notebook с комментариями о проделанной работе и полученными результатами.

Мини-проект

После выполнения практических заданий предлагается выбрать одно и выполнить его расширенную версию, или выбрать другую тему из тех, что обсуждались на семинарах, либо выбрать задачу и собрать данные самостоятельно. Цель – выполнить проект и подготовить отчет по нему в виде презентации или интерактивного дэшборда и отчитаться о проведенном исследовании перед «клиентом», описав задачи исследования, основную идею алгоритмов, которые использовались в работе, визуализируя результаты.

Критерии оценивания: объем проделанных исследований, выводы и рекомендации о дальнейших шагах по развитию проекта, умение понятно представить результаты и визуализировать данные, оценка проекта другими студентами

3 модуль

Презентация плана курсового проекта

Студенту необходимо представить план курсовой работы - индивидуального курсового проекта в форме мультимедийной презентации и устного выступления с последующими ответами на вопросы преподавателя и других студентов.

Содержание презентации должно включать в себя:

- техническое задание, цели и задачи проекта;
- краткий обзор актуальных практик, анализ рынка, ситуационный анализ;
- подробное описание этапов работы над проектом.

Более подробная структура презентации плана курсовой работы будет представлена преподавателем на первом занятии 3 модуля.

На этапе представления плана курсового проекта используются следующие критерии, которые в дальнейшем будут применяться также при оценивании результатов курсового проекта по итогам года руководителем курсовой работы и комиссией на защите:

Дескриптор	**Критерии оценивания**	
Студент выбирает методы, подходы и инструменты для решения поставленных профессиональных задач, исходя из критического анализа существующей бизнес-практики и обзора релевантных научных исследований.	*Неудовл.*	Студент не объясняет выбор методов, подходов и инструментов для решения поставленной задачи.
	Удовл.	Студент обзорно перечисляет некоторые используемые в бизнес-практике и исследованиях методы, подходы и инструменты, при этом не сопоставляет и не перерабатывает их, исходя из задач проекта.
	Хорошо	Студент описывает и сравнивает основные используемые в бизнес-практике и научных исследованиях методы, подходы и инструменты, однако не оценивает их критически, их выбор для задач проекта недостаточно обоснован.
	Отлично	Студент представляет исчерпывающий обзор релевантных методов, подходов и инструментов для решения профессиональной задачи, представляет их критический анализ, приводит подробное обоснование своего выбора, комбинирует и перерабатывает их, исходя из специфики проекта.
Студент разрабатывает поэтапный и развёрнутый план реализации проекта, исходя из технического задания; чётко ставит задачи и детализирует способы их реализации; понимает специфику и взаимосвязь бизнес-задач в области рекламы и связей с общественностью и задач анализа данных.	*Неудовл.*	План проекта непоследователен, способы реализации не прописаны, не соответствуют поставленным задачам.
	Удовл.	При разработке плана проекта студент не учитывает специфику задач в области рекламы и связей с общественностью, способы реализации прописаны недостаточно детально, нарушения логики этапов реализации проекта.
	Хорошо	План проекта логичен и соответствует техническому заданию, демонстрируется общее понимание специфики бизнес-задач и задач анализа данных, однако способы реализации задач прописаны недостаточно подробно.
	Отлично	Представлен детальный развёрнутый план реализации

		проекта, каждый этап логически следует за предыдущим и обоснован с учётом специфики бизнес-задач и задач анализа данных в соответствии с техническим заданием, подробно прописана методология и ресурсы реализации каждого этапа.
Студент демонстрирует навыки убедительного представления результатов профессиональной деятельности, презентация отличается выбором релевантных терминологических и визуальных средств	*Неудовл.*	Студент не владеет навыками убедительной презентации результатов профессиональной деятельности.
	Удовл.	Содержание курсового проекта не адаптировано для публичного представления, выбраны скудные терминологические и визуальные средства.
	Хорошо	Содержание курсового проекта адаптировано для публичного представления, студент делает акцент на основных результатах профессиональной деятельности, однако выступление недостаточно убедительно в части выбора релевантных терминологических и визуальных средств.
	Отлично	Содержание курсового проекта адаптировано для публичного представления, студент делает акцент на основных результатах профессиональной деятельности, выступление отличается выбором убедительных терминологических и визуальных средств.

Предзащита группового проекта.

Проектной группе необходимо представить результаты группового проекта в форме мультимедийной презентации и устного выступления с последующими ответами на вопросы преподавателя и других студентов. Презентация должна включать в себя описание бизнес-задачи и её решения с помощью методов анализа данных, наглядную визуализацию данных. Более подробная структура презентации будет представлена преподавателем.

Оценка предзащиты группового проекта – общая для проектной группы.

На этапе предзащиты групповых проектов используются критерии оценивания, которые в дальнейшем также будут применяться на защите проектов в формате экзамена по проектной деятельности в сессионный период по итогам 3 модуля:

Дескриптор	**Критерии оценивания**	
Студент владеет методами анализа данных, способен оценить качество данных и подготовить данные, построить гипотезы о нали-	*Неудовл.*	Студент не владеет методами анализа данных, не способен оценить качество данных, подготовить данные, построить гипотезы и выбрать релевантную методику моделирования.

чии закономерностей в наборах данных, выбрать методику моделирования, соответствующую профессиональной задаче в области коммуникаций.	*Удовл.*	Студент демонстрирует базовое владение некоторыми методами анализа данных, но не способен применить эти навыки для конкретного кейса; способен поверхностно описать данные, но не даёт исчерпывающий анализ их качества; гипотезы о наличии закономерностей в данных сформулированы недостаточно корректно; методика моделирования не соответствует профессиональной задаче в области коммуникаций.
	Хорошо	Студент демонстрирует владение основными методами анализа данных и применяет их для решения конкретной прикладной задачи; основные параметры оценки качества данных учтены; гипотезы сформулированы в целом корректно, однако не в полной мере соответствуют решаемой профессиональной задаче; выбранная методика моделирования подходит, но не оптимальна для проекта.
	Отлично	Студент демонстрирует продвинутый уровень владения методами анализа данных, выбирает и комбинирует их для решения профессиональной задачи; проводить всестороннюю оценку и предлагает решения проблем качества данных; гипотезы сформулированы корректно и соответствуют цели проекта; выбранная методика моделирования оптимальна.
Студент демонстрирует знания и профессиональную экспертизу в предметной области, логически обосновывает результаты проекта, прогнозы, рекомендации, вынесенные на защиту.	*Неудовл.*	Студент не владеет знаниями в предметной области, результаты проекта, прогнозы, рекомендации не обоснованы.
	Удовл.	Студент демонстрирует поверхностный уровень знаний в предметной области, не способен убедительно обосновать выносимые на защиту результаты, прогнозы, рекомендации, присутствуют нарушения логики.
	Хорошо	Студент демонстрирует базовые знания в предметной области, основные результаты проекта логически обоснованы, однако некоторые положения недостаточно аргументированы.
	Отлично	Студент демонстрирует высокий уровень экспертизы в предметной области, результаты проекта, рекомендации, прогнозы, выносимые на защиту, базируются на строгой логике и весомой аргументации.
Результаты проекта соответствуют целям и задачам, поставленным в	*Неудовл.*	Результаты проекта не соответствуют целям и задачам, поставленным в техническом задании; студент не даёт

техническом задании; студент способен давать развёрнутые ответы на вопросы и критические замечания по реализованному проекту.		содержательных ответов на вопросы и замечания по реализованному проекту.
	Удовл.	Результаты проекта не в полной мере соответствуют целям и задачам, поставленным в техническом задании; ответы на вопросы и замечания неполные.
	Хорошо	Результаты проекта соответствуют целям и задачам, поставленным в техническом задании, однако студент не аргументирует оптимальность решения по сравнению с другими альтернативами; студент даёт релевантные ответы на вопросы и замечания.
	Отлично	Результаты проекта в полной мере соответствуют целям и задачам, поставленным в техническом задании, студент аргументирует оптимальность решения по сравнению с возможными альтернативами; студент даёт исчерпывающие и убедительные ответы на вопросы и замечания.

4 модуль

Студенту необходимо представить презентацию результатов курсового проекта. Выступление на 10 минут, презентация от 10 слайдов, а также ответить на вопросы преподавателя по курсовому проекту. Критерии оценивания на предзащите соответствуют критериям качества защиты курсового проекта комиссией на защите.

Дескриптор	**Критерии оценивания**	
Студент демонстрирует знания и профессиональную экспертизу в предметной области, логически обосновывает результаты проекта, прогнозы, рекомендации, вынесенные на защиту.	*Неудовл.*	Студент не владеет знаниями в предметной области, результаты проекта, прогнозы, рекомендации не обоснованы.
	Удовл.	Студент демонстрирует поверхностный уровень знаний в предметной области, не способен убедительно обосновать выносимые на защиту результаты, прогнозы, рекомендации, присутствуют нарушения логики.
	Хорошо	Студент демонстрирует базовые знания в предметной области, основные результаты проекта логически обоснованы, однако некоторые положения недостаточно аргументированы.
	Отлично	Студент демонстрирует высокий уровень экспертизы в предметной области, результаты проекта, рекомендации, прогнозы, выносимые на защиту, базируются на строгой логике и весомой аргументации.
Результаты проекта соответствуют целям и	*Неудовл.*	Результаты проекта не соответствуют целям и задачам, поставленным в техническом задании; студент не даёт содержательных ответов на

задачам, поставленным в техническом задании; студент способен давать развёрнутые ответы на вопросы и критические замечания по реализованному проекту.		вопросы и замечания комиссии по реализованному проекту.
	Удовл.	Результаты проекта не в полной мере соответствуют целям и задачам, поставленным в техническом задании; ответы на вопросы и замечания неполные.
	Хорошо	Результаты проекта соответствуют целям и задачам, поставленным в техническом задании, однако студент не аргументирует оптимальность решения по сравнению с другими альтернативами; студент даёт релевантные ответы на вопросы и замечания комиссии.
	Отлично	Результаты проекта в полной мере соответствуют целям и задачам, поставленным в техническом задании, студент аргументирует оптимальность решения по сравнению с возможными альтернативами; студент даёт исчерпывающие и убедительные ответы на вопросы и замечания комиссии.
Студент демонстрирует навыки убедительного представления результатов профессиональной деятельности, презентация отличается выбором релевантных терминологических и визуальных средств	*Неудовл.*	Студент не владеет навыками убедительной презентации результатов профессиональной деятельности.
	Удовл.	Содержание курсового проекта не адаптировано для публичного представления, выбраны скудные терминологические и визуальные средства.
	Хорошо	Содержание курсового проекта адаптировано для публичного представления, студент делает акцент на основных результатах профессиональной деятельности, однако выступление недостаточно убедительно в части выбора релевантных терминологических и визуальных средств.
	Отлично	Содержание курсового проекта адаптировано для публичного представления, студент делает акцент на основных результатах профессиональной деятельности, выступление отличается выбором убедительных терминологических и визуальных средств.

Экзамен

На устном экзамене оцениваются знания студента в соответствии с программой курса.

8. Содержание дисциплины

2 модуль

Экосистема рекламного (digital) рынка и его эволюция.

Как устроен рекламный рынок и какие игроки его представляют, взаимодействие между участниками.

Индикаторы эффективной коммуникации.

По каким метрикам бренды (компании) оценивают эффективность медиа и коммуникационных активностей.

Инструменты измерения эффективности.
Обзор современных инструментов измерения эффективности маркетинговых активностей бренда.

Инструменты увеличения эффективности.
Обзор современных инструментов и подходов, которые помогают рекламодателям увеличить эффективность коммуникации.

Онлайн данные.
Какие данные существуют, как собрать данные, как их использовать?

Оффлайн данные.
Какие данные существуют, как собрать данные, как их использовать?

Онлайн + Оффлайн.
Как связывают ранее несвязные сущности? Обзор подходов и методов. Успешные примеры

Смежные источники данных и аналитики.
Мониторинг открытых источников и данные социальных сетей.

От медиа планирования к аудиторному планированию.
Изменение подходов к планированию рекламных активностей бренда.

Безопасность данных.
Безопасность данных, ограничения со стороны государства и влияние мировых рынков.

От обезличенной коммуникации к персональной.
Решение задач кросс-девайс коммуникации и задач омниканальности.

Изменение потребительского поведения.
Как меняется современное потребление, и какие технологии будут актуальны через 5 лет.

Анализ данных для решения задач рекламодателей.
Ключевые концепты, основные тренды, AdTech, Data Market, основные проблемы. Обзор инструментов Data science и применения в маркетинге. Полезные ресурсы.

Первичный анализ данных.
Выгрузка и предобработка данных (GCP, bash), демонстрация основных методов Pandas, первые попытки описания аудитории, статистические сравнения групп пользователей. Демонстрация основных методов matplotlib, seaborn, plotly.

Задачи обучения с учителем. Как построить свой Look-a-like.
Алгоритмы классификации. Деревья решений и метод ближайших соседей. Как строится, работает, параметры и применение в реальных задачах. Линейные модели классификации и регрессии. Random Forest. Оценка эффективности алгоритмов. Построение скоринговых моделей для кук.

Проект 1.

Обучение без учителя. Поиск групп среди аудитории сайта и построение персональной коммуникации.
Метод главных компонент и кластеризация. Особенности применения PCA, выбор признаков, методы и метрики качества. Проект 2.

Другие прикладные задачи анализа данных.
Особенности задач, методы построения выводов, на что обратить внимание. Кредитный скоринг, задача оттока и анализ эффективности удержания, анализ результатов АБ теста. Кейсы.

Оценка эффективности рекламных кампаний.
Комплексная оценка метрик, моделирование атрибуции, эвристические и мультиканальные модели, их реализация и оценка эффективности РК.

Визуализация и представление результатов.
Как визуализировать отчет так, чтобы понравилось клиенту. Построение интерактивных дэшбордов. Обзор BI систем. Google Studio, Bime, Tableau.

3 модуль

В 3 модуле проектно-исследовательский семинар выполняет основную функцию поддержки всех форм проектной и научно-исследовательской деятельности студентов в рамках образовательной программы:

- подготовка курсовых работ в формате индивидуальных курсовых проектов;
- выполнение групповых проектов.

Введение в проектную работу.
Особенности реализации коммуникационных проектов, основанных на данных: бизнес-практика, кейсы. Обсуждение выбранных студентами тем курсовых работ и групповых проектов. Требования к курсовым проектам и групповой проектной работе. Согласование графика презентаций планов курсовых работ и предзащиты групповых проектов.

Построение и проверка гипотез в интеллектуальном анализе данных.
Эмпирически верифицируемая гипотеза. Жизненный цикл гипотезы в интеллектуальном анализе данных. Построение гипотез. Особенности проверки гипотез при работе с большими данными. Кросс-валидация. Обучающая и контрольная выборка. Уточнение гипотез.

Презентация и обсуждение планов курсовых работ.
Презентация каждым студентом плана индивидуального курсового проекта в соответствии с утверждённым графиком. Ответы на вопросы преподавателя и других студентов. Групповое обсуждение, рекомендации преподавателя по работе над проектом.

Предзащита и обсуждение групповых проектов.
Предзащита группового проекта в соответствии с утверждённым графиком. Ответы на вопросы преподавателя и других студентов. Групповое обсуждение, рекомендации преподавателя по доработке проекта.

4 модуль

Интеллектуальный анализ текста

Предварительная обработка текстовых данных: токены, N-граммы. Тематическое моделирование. Извлечение информации и кластеризация. Байесовская классификация текстов. Анализ тональности текстов. Глубокое обучение в интеллектуальном анализе текста. Применение интеллектуального анализа текстов для прикладных исследований в области рекламы и связей с общественностью (анализ потребительских предпочтений, оценка репутации, мониторинг трендов и др.).

Введение в сетевой анализ
Основные понятия теории графов и сетевого анализа. Типы графовых моделей социальной сети. Измеримые параметры социальной сети: плотность, центральность, распределение степеней и др. Анализ и интерпретация социальных графов. Дизайн исследования в сетевом анализе. Применение сетевого анализа для прикладных исследований в области рекламы и связей с общественностью (поиск агентов влияния для распространения вирального контента, анализ негативного информационного поля в социальных медиа и др.).

Предзащита курсовых работ

9. Образовательные технологии

Обзор практических задач и кейсов. Встречи с представителями коммуникационной индустрии, мастер-классы экспертов. Работа с ПО для анализа и визуализации данных.

10. Оценочные средства для текущего контроля и аттестации студента

Примеры заданий промежуточной аттестации по 2 модулю курса.

1. Анализ и сравнение аудиторных групп пользователей – чем покупатели отличаются от случайных посетителей?

2. Поиск людей, похожих на тех, кто совершил покупку, расширение сегмента потенциальных покупателей с помощью скоринга.

3. Сегментация пользователей методами кластерного анализа. Выбор эффективного кластера.

Примерные вопросы для устного экзамена.

1. Что такое Data, какие бывают типы данных: 1st part, 2nd party, 3rd party data.
2. Что такое DMP, основные функции.
3. Как обеспечивается защита данных.
4. Поставщики данных.
5. Крупные игроки на рынке data-продуктов.
6. Основные тренды развития рынка DATA .
7. Пример кейса Data-driven подхода с DMP (задача - методы решения - результат)
8. Инструменты измерения и увеличения эффективности.
9. Подходы к планированию медиа активности бренда.
10. Решение задач кросс-девайс коммуникации и задач омниканальности
11. Задачи анализа данных в рекламе.
12. Основные тренды развития рынка DATA на европейском и российском рынке.
13. Методы интеллектуального анализа данных.

14. Построение и проверка гипотез при работе с большими данными.
15. Измеримые параметры социальной сети в сетевом анализе, анализ и интерпретация социальных графов.
16. Предварительная обработка текстовых данных.
17. Байесовская классификация текстов.
18. Анализ тональности текстов и его применение в коммуникационных исследованиях.

11. Порядок формирования оценок по дисциплине

В диплом выставляется результирующая оценка по учебной дисциплине:

$$O_{результ} = 0,7 * O_{накопл} + 0,3 * \cdot O_{экз}$$

Где:

$\boldsymbol{O_{экз}}$ – оценка за итоговый контроль в форме устного экзамена.

Устный экзамен по билетам. В билете по 2 вопроса, имеющих одинаковый вес. Каждый вопрос оценивается по 10-балльной шкале. Округление: арифметическое. Оценки до 4 баллов не округляются в большую сторону – знаки после запятой отсекаются. Например, оценка 7,76 округляется до 8 баллов, оценка 3,81 округляется до 3 баллов.

Накопленная оценка рассчитывается по формуле:

$$O_{накопл} = 0,4 * O_{промежуточная1} + 0,3 * O_{промежуточная2} + 0,3 * O_{промежуточная3}$$

где:

$\mathbf{O}_{промежуточная1}$ - оценка за второй модуль

$\mathbf{O}_{промежуточная2}$ - оценка за третий модуль

$\mathbf{O}_{промежуточная3}$ - оценка за четвёртый модуль

Округление промежуточных оценок и результирующей оценки – арифметическое. Оценки до 4 баллов не округляются в большую сторону – знаки после запятой отсекаются. Например, оценка 7,76 округляется до 8 баллов, оценка 3,81 округляется до 3 баллов.

Промежуточная оценка за второй модуль выставляется по следующей формуле:

$$O_{промежуточная1} = 0,2 * O_{ауд.} + 0,5 * O_{проект} + 0,1 * O_{дз1} + 0,1 * O_{дз2} + 0,1 * O_{дз3}$$

$O_{ауд}$ – оценка за активность и качество аудиторной работы.

Преподаватель оценивает работу студентов на семинарских занятиях: активность участия в дискуссиях, правильность решения задач и практических кейсов. Оценка по 10-балльной шкале выставляется за работу на каждом семинаре, суммируется и делится на количество семинаров. В случае отсутствия студента на семинаре без уважительной причины (подтверждается справкой), за семинар выставляется 0 баллов. Округление аудиторной оценки: арифметическое.

$O_{проект}$ – оценка за итоговый мини-проект второго модуля.

$O_{дз1}, O_{дз2}, O_{дз3}$ – оценки за домашние практические задания по разделам второго модуля.

Оценки за все формы текущего контроля выставляются по 10-балльной шкале.

Промежуточная оценка за третий модуль выставляется по следующей формуле:

$O_{промежуточная2} = 0{,}4*O_{ауд.} + 0{,}3*O_{план} + 0{,}3 * O_{предзащита}$

$O_{ауд}$ – оценка за активность и качество аудиторной работы.

Преподаватель оценивает решение практических задач по разделу «Построение и проверка гипотез в интеллектуальном анализе данных» на семинарах. Оценка по 10-балльной шкале выставляется за выполнение каждого задания. Невыполненное задание оценивается в 0 баллов.

Преподаватель оценивает участие каждого студента в групповых дискуссиях. Каждый студент принимает участие в обсуждении планов курсовых проектов других студентов и предзащиты проектов других групп.

Оценивается качество вопросов и комментариев по следующим критериям:

- уровень подготовки к теме обсуждения, демонстрация знаний в предметной области,
- релевантность вопросов, исходя из содержания докладов коллег, критическое мышление,
- ценность вопросов и комментариев, исходя из задач и специфики проекта.

Оценка по 10-балльной шкале выставляется за обсуждение каждой презентации других студентов. Студент, присутствовавший на презентации, но не принимавший участие в дискуссии, получает оценку от 1 до 3 баллов. В случае отсутствия студента без уважительной причины (подтверждается справкой), за участие в дискуссии выставляется 0 баллов.

Вес выполненных заданий и участия в обсуждении проектов равный. По итогам модуля все оценки суммируются и делятся на количество активностей.

Округление аудиторной оценки: арифметическое.

$O_{план}$ – оценка за презентацию плана индивидуального курсового проекта.

$O_{предзащита}$ – оценка за предзащиту группового проекта.

Оценки за все формы текущего контроля выставляются по 10-балльной шкале.

Промежуточная оценка за четвёртый модуль выставляется по следующей формуле:

$O_{промежуточная3} = 0{,}5*O_{ауд.} + 0{,}5*O_{предзащита2}$

$O_{ауд}$ – оценка за активность и качество аудиторной работы.

Преподаватель оценивает решение практических задач и участие в дискуссиях на каждом семинаре. Оценка по 10-балльной шкале выставляется за работу на каждом семинаре, суммируется и делится на количество семинаров. В случае отсутствия студента на семинаре без уважительной причины (подтверждается справкой), за семинар выставляется 0 баллов. Округление аудиторной оценки: арифметическое.

$O_{предзащита2}$ – оценка за предзащиту курсового проекта.

12. Учебно-методическое и информационное обеспечение дисциплины

Базовый учебник

Hastie, T., Tibshirani, R., & Friedman, J. (2009). The Elements of Statistical Learning, 2nd edition. Springer.

Электронная версия базового учебника доступна в библиотеке НИУ ВШЭ: https://proxylibrary.hse.ru:2104/book/10.1007/978-0-387-84858-7.

Дополнительная литература

Воронцов К.В. Алгоритмы кластеризации и многомерного шкалирования. Курс лекций. М: МГУ, 2007.

Foster, I., Ghani, R., Jarmin, R. S., Kreuter, F., Lane, J. (2017). Big Data and Social Science: A Practical Guide to Methods and Tools. CRC Press.

Ishikawa, H. (2015). Social Big Data Mining. CRC Press.

Kelleher, J. D. (2015). Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies. The MIT Press.

Lee, H. & Sohn, I. (2016). Fundamentals of Big Data Network Analysis for Research and Industry. Wiley.

Powell, V. Principal Component Analysis Explained Visually. URL: http://setosa.io/ev/principal-component-analysis/.

Thai, M., Wu, W., Xiong, H. (2017). Big Data in Complex and Social Networks. CRC Press.

Справочники, словари, энциклопедии

Документация sklearn Model evaluation: quantifying the quality of predictions. URL: http://scikit-learn.org/stable/modules/model_evaluation.html.

Программные средства

Python, пакет Anaconda для анализа данных.
R, R Studio (опционально)
ПО для визуализации данных: Google Studio, Bime, Tableau.

Дистанционная поддержка дисциплины

Электронная почта

13. Материально-техническое обеспечение дисциплины

Компьютер, проектор, компьютерный класс для выполнения практических заданий. Рекомендуется использование ноутбука.