

**МИНИСТЕРСТВО ЦИФРОВОГО РАЗВИТИЯ,
СВЯЗИ И МАССОВЫХ КОММУНИКАЦИЙ РОССИЙСКОЙ ФЕДЕРАЦИИ**
**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ**
**«САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ТЕЛЕКОММУНИКАЦИЙ ИМ. ПРОФ. М.А. БОНЧ-БРУЕВИЧА»**
(СПбГУТ)

Кафедра _____ Бизнес-информатики _____
(полное наименование кафедры)



Регистрационный №_21.07/6-Д

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Анализ данных

(наименование дисциплины)

образовательная программа высшего образования

38.03.05 Бизнес-информатика

(код и наименование направления подготовки / специальности)

бакалавр

(квалификация)

Управление бизнес-процессами предприятия в цифровой экономике

(направленность / профиль образовательной программы)

очная форма

(форма обучения)

Санкт-Петербург

Рабочая программа дисциплины составлена на основе требований Федерального государственного образовательного стандарта высшего образования по направлению (специальности) подготовки «38.03.05 Бизнес-информатика», утвержденного приказом Министерства образования и науки Российской Федерации от 29.07.2020 № 838, и в соответствии с рабочим учебным планом, утвержденным ректором университета.

1. Цели и задачи дисциплины

Целью преподавания дисциплины «Анализ данных» является:
формирование у студентов системного представления о методах и технологиях интеллектуального анализа данных

Эта цель достигается путем решения следующих(ей) задач(и):

дать представление о моделях архитектуры хранилищ данных; познакомить с методами Data Mining и сферами их применения; научить применять технологии интеллектуального анализа данных для решения организационно-экономических задач

2. Место дисциплины в структуре образовательной программы

Дисциплина «Анализ данных» Б1.В.03 является дисциплиной части, формируемой участниками образовательных отношений блока 1 учебного плана подготовки бакалавриата по направлению «38.03.05 Бизнес-информатика». Изучение дисциплины «Анализ данных» опирается на знания дисциплин(ы) «Информационное обеспечение бизнес-процессов предприятия»; «Теория вероятностей и прикладная статистика».

3. Перечень планируемых результатов обучения, соотнесенных с планируемыми результатами освоения образовательной программы

Процесс изучения дисциплины направлен на формирование следующих компетенций:

Таблица 1

№ п/п	Код компетенции	Наименование компетенции
1	ПК-12	Способен проводить интеллектуальный анализ данных для принятия обоснованных организационно-управленческих решений

Индикаторы достижения компетенций

Таблица 2

ПК-12.1	Знать: особенности и правила сбора и обработки статистических данных; методы и технологии интеллектуального анализа данных; принципы работы с большими данными; задачи и области применения статистических методов и методов интеллектуального анализа данных; методы и технологии визуализации деловой информации; область применения различных видов статистических показателей; методы анализа и оценки достоверности собранной информации
ПК-12.2	Уметь: использовать статистические методы и методы интеллектуального анализа данных для принятия обоснованных организационно-управленческих решений; обрабатывать и оформлять результаты аналитической деятельности; оценивать достоверность полученной информации; подготавливать презентации, использовать диаграммы, инфографику и информационные панели для визуализации результатов анализа

ПК-12.3	Владеть: методами сбора, обработки и анализа статистических данных; навыками анализа собранной информации на соответствие установленным критериям; технологиями интеллектуального анализа данных для решения организационно-экономических задач; методами количественного анализа и моделирования; инструментальными средствами поддержки принятия управленческих решений; инструментальными средствами визуализации деловой информации
---------	---

4. Объем дисциплины и виды учебной работы

Очная форма обучения

Таблица 3

Вид учебной работы		Всего часов	Семестры
			5
Общая трудоемкость	4 ЗЕТ	144	144
Контактная работа с обучающимися		52.35	52.35
в том числе:			
Лекции		20	20
Практические занятия (ПЗ)		16	16
Лабораторные работы (ЛР)		14	14
Защита контрольной работы			-
Защита курсовой работы			-
Защита курсового проекта			-
Промежуточная аттестация		2.35	2.35
Самостоятельная работа обучающихся (СРС)		58	58
в том числе:			
Курсовая работа			-
Курсовой проект			-
И / или другие виды самостоятельной работы: подготовка к лабораторным работам, практическим занятиям, контрольным работам, изучение теоретического материала		58	58
Подготовка к промежуточной аттестации		33.65	33.65
Вид промежуточной аттестации			Экзамен

5. Содержание дисциплины

5.1. Содержание разделов дисциплины.

Таблица 4

№ п/п	Наименование раздела (темы) дисциплины	Содержание раздела	№ семестра		
			очная	очно-заочная	заочная
1	Раздел 1. Системы поддержки принятия решений	Big data. Data Science. Системы оперативной обработки информации. Системы поддержки принятия решений. Business Intelligence.	5		
2	Раздел 2. Хранилища данных	Хранилище данных. ETL. Архитектура хранилищ данных. Витрины данных. Озера данных	5		

3	Раздел 3. Оперативный анализ данных	Технология OLAP (Online Analytical Processing). Работа с измерениями. Data Mining. Искусственный интеллект. Машинное обучение	5		
4	Раздел 4. Основные задачи интеллектуального анализа данных	Классификация данных. Кластеризация данных. Прогнозирование данных. Поиск ассоциативных правил	5		
5	Раздел 5. Основные методы и алгоритмы Data Mining	Корреляционно-регрессионный анализ. Логистическая регрессия. Деревья решений. Нейронные сети. Карта Кохонена. Алгоритм EM. Алгоритм Clope. Алгоритм k-means (k-средних). Алгоритм Apriori. Алгоритм FPG. Визуализация данных. Web mining. Process Mining	5		

5.2. Междисциплинарные связи с обеспечиваемыми (последующими) дисциплинами.

Таблица 5

№ п/п	Наименование обеспечиваемых (последующих) дисциплин
1	Маркетинг
2	Финансовый менеджмент

5.3. Разделы дисциплин и виды занятий.

Очная форма обучения

Таблица 6

№ п/п	Наименование раздела (темы) дисциплин	Лек-ции	Практ. занятия	Лаб. занятия	Семи-нары	СРС	Всего часов
1	Раздел 1. Системы поддержки принятия решений	2				10	12
2	Раздел 2. Хранилища данных	4				10	14
3	Раздел 3. Оперативный анализ данных	2	4			10	16
4	Раздел 4. Основные задачи интеллектуального анализа данных	4		2		10	16
5	Раздел 5. Основные методы и алгоритмы Data Mining	8	12	12		18	50
Итого:		20	16	14	-	58	108

6. Лабораторный практикум

Очная форма обучения

Таблица 7

№ п/п	Номер раздела (темы)	Наименование лабораторной работы	Всего часов
1	4	Прогнозирование временных рядов	2
2	5	Корреляционный анализ	2
3	5	Регрессионный анализ	2
4	5	Логистическая регрессия	2

5	5	Предсказательная аналитика с помощью нейронной сети	4
6	5	Кластеризация данных	2
Итого:			14

7. Практические занятия (семинары)

Очная форма обучения

Таблица 8

№ п/п	Номер раздела (темы)	Наименование практических занятий (семинаров)	Всего часов
1	3	Работа с OLAP кубом	2
2	3	Знакомство с аналитической платформой Loginom. Импорт и предобработка данных	2
3	5	RFM-анализ	2
4	5	Обработка результатов опроса потребителей	2
5	5	Анализ транзакций	2
6	5	Анализ динамики структуры чека	2
7	5	Расчет скорингового балла	2
8	5	ABC-XYZ-анализ	2
Итого:			16

8. Примерная тематика курсовых проектов (работ)

Рабочим учебным планом не предусмотрено

9. Самостоятельная работа

Очная форма обучения

Таблица 9

№ раздела дисциплины	Содержание СРС	Форма контроля	Всего часов
1	Изучение теоретического материала	опрос	10
2	Изучение теоретического материала	опрос	10
3	Изучение теоретического материала. Подготовка к практическим занятиям	опрос	10
4	Изучение теоретического материала	опрос	10
5	Изучение теоретического материала. Подготовка к практическим и лабораторным занятиям	опрос	18
Итого:			58

10. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине

Для самостоятельной работы по дисциплине рекомендовано следующее учебно-методическое обеспечение:

- Положение о самостоятельной работе студентов в Санкт-Петербургском

- государственном университете телекоммуникаций им. проф. М.А. Бонч-Бруевича;
- рекомендованная основная и дополнительная литература;
 - конспект занятий по дисциплине;
 - слайды-презентации и другой методический материал, используемый на занятиях;
 - методические рекомендации по подготовке письменных работ, требования к их содержанию и оформлению (реферат, эссе, контрольная работа) ;
 - фонды оценочных средств;
 - методические указания к выполнению лабораторных работ для студентов;

11. Фонд оценочных средств для проведения промежуточной аттестации обучающихся

Фонд оценочных средств разрабатывается в соответствии с локальным актом университета "Положение о фонде оценочных средств" и является приложением (Приложение А) к рабочей программе дисциплины.

Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине включает в себя:

- перечень компетенций с указанием этапов их формирования в процессе освоения образовательной программы;
- описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания;
- типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы;
- методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций.

Для каждого результата обучения по дисциплине определяются показатели и критерии оценки сформированности компетенций на различных этапах их формирования, шкалы и процедуры оценивания.

12. Перечень основной и дополнительной литературы, необходимой для освоения дисциплины

12.1. Основная литература:

1. Вольфсон, Михаил Борисович.
Анализ данных : [Электронный ресурс] : учеб. пособие / М. Б. Вольфсон ; рец.: Ю. П. Левчук, А. Л. Алимов ; Федер. агентство связи, С.-Петербург. гос. ун-т телекоммуникаций им. проф. М. А. Бонч-Бруевича. - СПб. : СПбГУТ, 2015. - 81 с. : ил. - Библиогр.: с. 81. - 451.36 р.
2. Чубукова, И. А.
Data Mining : [Электронный ресурс] : учебное пособие / И. А. Чубукова. - 2-е изд. - Москва : ИНТУИТ, 2016. - 470 с. - URL: <https://e.lanbook.com/book/100582>. - ISBN 978-5-94774-819-2 : Б. ц. Книга из коллекции ИНТУИТ - Информатика

12.2. Дополнительная литература:

1. Вольфсон, Михаил Борисович.
Анализ данных : [Электронный ресурс] : методические указания к лабораторным работам / М. Б. Вольфсон ; рец. Левчук ; Федеральное агентство связи, С.-Петербург. гос. ун-т телекоммуникаций им. проф. М. А. Бонч-Бруевича. - СПб. : СПбГУТ, 2013. - 47 с. : ил. - .
2. Анализ данных и процессов. 3-е изд. : [Электронный ресурс] / А. Барсегян, М. Куприянов, И. Холод и др. - СПб. : БХВ-Петербург, 2010. - 512 с. : ил. - URL: <http://ibooks.ru/reading.php?productid=18456>. - ISBN 978-5-9775-0368-6 : Б. ц.
3. Андрианова, Екатерина Евгеньевна.
Управление данными. Интеллектуальный анализ данных : [Электронный ресурс] : учебное пособие / Е. Е. Андрианова, И. А. Липанова, О. Ю. Сабинин ; рец.: О. Ю. Ильяшенко, Е. В. Давыдова ; Федер. агентство связи, Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования "Санкт-Петербургский государственный университет телекоммуникаций им. проф. М. А. Бонч-Бруевича". - СПб. : СПбГУТ, 2016. - 38 с. : ил. - 214.93 р.
4. Барский, А. Б.
Введение в нейронные сети : [Электронный ресурс] : учебное пособие / А. Б. Барский. - 2-е изд. - Москва : ИНТУИТ, 2016. - 358 с. - URL: <https://e.lanbook.com/book/100684>. - Б. ц. Книга из коллекции ИНТУИТ - Информатика
5. Жуковский, О. И.
Информационные технологии и анализ данных : [Электронный ресурс] : учебное пособие / О. И. Жуковский. - Москва : ТУСУР, 2014. - 130 с. - URL: <https://e.lanbook.com/book/110351>. - ISBN 978-5-4332-0158-3 : Б. ц. Книга из коллекции ТУСУР - Информатика

13. Перечень ресурсов информационно-телекоммуникационной сети «Интернет»

Ресурсы информационно-телекоммуникационной сети «Интернет» из указанного перечня являются рекомендуемыми дополнительными (вспомогательными) источниками официальной информации, размещенной на легальных основаниях с открытым доступом. За полноту содержания и качество работу сайтов несет ответственность правообладатель.

Таблица 10

Наименование ресурса	Адрес
BaseGroup Labs (Документация)	basegroup.ru/deductor/manual
BaseGroup Labs (Статьи)	basegroup.ru/community/articles
НОУ Интуит «Data Mining»	www.intuit.ru/studies/courses/6/6/info
НОУ Интуит «Проектирование хранилищ данных для приложений систем деловой осведомленности (Business Intelligence Systems)»	www.intuit.ru/studies/courses/599/455/info

14. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень программного обеспечения и информационных справочных систем.

14.1. Программное обеспечение дисциплины:

- Deductor Studio Academic
- Google Chrome
- Open Office
- Windows 7 ЦЭУБИ

14.2. Информационно-справочные системы:

- ЭБС iBooks (<https://ibooks.ru>)
- ЭБС СПбГУТ (<http://lib.spbgut.ru>)
- ЭБС Лань (<https://e.lanbook.com/>)

15. Методические указания для обучающихся по освоению дисциплины

15.1. Планирование и организация времени, необходимого для изучения дисциплины

Важным условием успешного освоения дисциплины «Анализ данных» является создание системы правильной организации труда, позволяющей распределить учебную нагрузку равномерно в соответствии с графиком образовательного процесса. Большую помощь в этом может оказать составление плана работы на семестр, месяц, неделю, день. Его наличие позволит подчинить свободное время целям учебы, трудиться более успешно и эффективно. Нужно осуществлять самоконтроль, который является необходимым условием успешной учебы. Все задания, включая вынесенные на самостоятельную работу, рекомендуется выполнять непосредственно после соответствующего аудиторного занятия (лекции, практического занятия), что способствует лучшему усвоению материала, позволяет своевременно выявить и устранить «пробелы» в знаниях, систематизировать ранее пройденный материал, на его основе приступить к овладению новыми знаниями и навыками.

Система университетского обучения основывается на рациональном сочетании нескольких видов учебных занятий (в первую очередь, лекций и практических занятий), работа на которых обладает определенной спецификой.

15.2. Подготовка к лекциям

Знакомство с дисциплиной происходит уже на первой лекции, где от студента требуется не просто внимание, но и самостоятельное оформление конспекта. При работе с конспектом лекций необходимо учитывать тот фактор, что одни лекции дают ответы на конкретные вопросы темы, другие – лишь выявляют взаимосвязи между явлениями, помогая студенту понять глубинные процессы развития изучаемого предмета, как в истории, так и в настоящее время.

Конспектирование лекций – сложный вид вузовской аудиторной работы, предполагающий интенсивную умственную деятельность студента. Конспект является полезным тогда, когда записано самое существенное и сделано это самим обучающимся. Не надо стремиться записать дословно всю лекцию. Такое «конспектирование» приносит больше вреда, чем пользы. Целесообразно вначале понять основную мысль, излагаемую лектором, а затем записать ее. Желательно запись осуществлять на одной странице листа или оставляя поля, на которых позднее, при самостоятельной работе с конспектом, можно сделать дополнительные записи, отметить непонятные места.

Конспект лекции лучше подразделять на пункты, соблюдая красную строку. Этому в большой степени будут способствовать вопросы плана лекции, предложенные преподавателям. Следует обращать внимание на акценты, выводы, которые делает лектор, отмечая наиболее важные моменты в лекционном материале замечаниями «важно», «хорошо запомнить» и т.п. Можно делать это и с помощью разноцветных маркеров или ручек, подчеркивая термины и определения.

Целесообразно разработать собственную систему сокращений, аббревиатур и символов. Однако при дальнейшей работе с конспектом символы лучше заменить обычными словами для быстрого зрительного восприятия текста. Работая над конспектом лекций, всегда необходимо использовать не только учебник, но и ту литературу, которую дополнительно рекомендовал лектор. Именно такая серьезная, кропотливая работа с лекционным материалом позволит глубоко овладеть теоретическим материалом.

15.3. Подготовка к практическим занятиям

Тщательное продумывание и изучение вопросов плана основывается на проработке пройденного материала (материала лекций, практических занятий), а затем изучения обязательной и дополнительной литературы, рекомендованной к данной теме.

Результат такой работы должен проявиться в способности студента свободно ответить на теоретические вопросы практикума, его выступлении и участии в коллективном обсуждении вопросов изучаемой темы, правильном выполнении практических заданий и контрольных работ.

Необходимо понимать, что невозможно во время аудиторных занятий изложить весь материал из-за лимита аудиторных часов, и при изучении дисциплины недостаточно конспектов занятий. Поэтому самостоятельная работа с учебниками, учебными пособиями, научной, справочной литературой, материалами периодических изданий и Интернета является наиболее эффективным методом получения дополнительных знаний, позволяет значительно активизировать процесс овладения информацией, способствует более глубокому усвоению изучаемого материала, формирует у студентов свое отношение к конкретной проблеме.

15.4. Рекомендации по работе с литературой

Работу с литературой целесообразно начать с изучения общих работ по теме, а также учебников и учебных пособий. Далее рекомендуется перейти к анализу монографий и статей, рассматривающих отдельные аспекты проблем, изучаемых в рамках курса, а также официальных материалов и неопубликованных документов

(научно-исследовательские работы, диссертации), в которых могут содержаться основные вопросы изучаемой проблемы.

Работу с источниками надо начинать с ознакомительного чтения, т.е. просмотреть текст, выделяя его структурные единицы. При ознакомительном чтении закладками отмечаются те страницы, которые требуют более внимательного изучения. В зависимости от результатов ознакомительного чтения выбирается дальнейший способ работы с источником. Если для разрешения поставленной задачи требуется изучение некоторых фрагментов текста, то используется метод выборочного чтения. Если в книге нет подробного оглавления, следует обратить внимание ученика на предметные и именные указатели.

Избранные фрагменты или весь текст (если он целиком имеет отношение к теме) требуют вдумчивого, неторопливого чтения с «мысленной проработкой» материала. Такое чтение предполагает выделение: 1) главного в тексте; 2) основных аргументов; 3) выводов. Особое внимание следует обратить на то, вытекает тезис из аргументов или нет. Необходимо также проанализировать, какие из утверждений автора носят проблематичный, гипотетический характер и уловить скрытые вопросы.

Понятно, что умение таким образом работать с текстом приходит далеко не сразу. Наилучший способ научиться выделять главное в тексте, улавливать проблематичный характер утверждений, давать оценку авторской позиции – это сравнительное чтение, в ходе которого студент знакомится с различными мнениями по одному и тому же вопросу, сравнивает весомость и доказательность аргументов сторон и делает вывод о наибольшей убедительности той или иной позиции.

Если в литературе встречаются разные точки зрения по тому или иному вопросу из-за сложности прошедших событий и правовых явлений, нельзя их отвергать, не разобравшись. При наличии расхождений между авторами необходимо найти рациональное зерно у каждого из них, что позволит глубже усвоить предмет изучения и более критично оценивать изучаемые вопросы. Знакомясь с особыми позициями авторов, нужно определять их схожие суждения, аргументы, выводы, а затем сравнивать их между собой и применять из них ту, которая более убедительна.

Следующим этапом работы с литературными источниками является создание конспектов, фиксирующих основные тезисы и аргументы. Можно делать записи на отдельных листах, которые потом легко систематизировать по отдельным темам изучаемого курса. Другой способ – это ведение тематических тетрадей-конспектов по одной какой-либо теме. Большие специальные работы монографического характера целесообразно конспектировать в отдельных тетрадях. Здесь важно вспомнить, что конспекты пишутся на одной стороне листа, с полями и достаточным для исправления и ремарок межстрочным расстоянием (эти правила соблюдаются для удобства редактирования). Если в конспектах приводятся цитаты, то непременно должно быть дано указание на источник (автор, название, выходные данные, № страницы). Впоследствии эта информация может быть использована при написании текста реферата или другого задания.

Таким образом, при работе с источниками и литературой важно уметь:

- сопоставлять, сравнивать, классифицировать, группировать, систематизировать информацию в соответствии с определенной учебной задачей;

- обобщать полученную информацию, оценивать прослушанное и прочитанное;
- фиксировать основное содержание сообщений; формулировать, устно и письменно, основную идею сообщения; составлять план, формулировать тезисы;
- готовить и презентовать развернутые сообщения типа доклада;
- работать в разных режимах (индивидуально, в паре, в группе), взаимодействуя друг с другом;
- пользоваться реферативными и справочными материалами;
- контролировать свои действия и действия своих товарищей, объективно оценивать свои действия;
- обращаться за помощью, дополнительными разъяснениями к преподавателю, другим студентам;
- пользоваться лингвистической или контекстуальной догадкой, словарями различного характера, различного рода подсказками, опорами в тексте (ключевые слова, структура текста, предваряющая информация и др.);
- использовать при говорении и письме перифраз, синонимичные средства, слова-описания общих понятий, разъяснения, примеры, толкования, «словотворчество»
- повторять или перефразировать реплику собеседника в подтверждении понимания его высказывания или вопроса;
- обратиться за помощью к собеседнику (уточнить вопрос, переспросить и др.);
- использовать мимику, жесты (вообще и в тех случаях, когда языковых средств не хватает для выражения тех или иных коммуникативных намерений).

15.5. Подготовка к промежуточной аттестации

При подготовке к промежуточной аттестации целесообразно:

- внимательно изучить перечень вопросов и определить, в каких источниках находятся сведения, необходимые для ответа на них;
- внимательно прочитать рекомендованную литературу;
- составить краткие конспекты ответов (планы ответов).

16. Материально-техническое обеспечение дисциплины

Таблица 11

№ п/п	Наименование специализированных аудиторий и лабораторий	Наименование оборудования
1	Лекционная аудитория	Аудио-видео комплекс
2	Аудитории для проведения групповых и практических занятий	Аудио-видео комплекс
3	Компьютерный класс	Персональные компьютеры
4	Аудитория для курсового и дипломного проектирования	Персональные компьютеры
5	Аудитория для самостоятельной работы	Компьютерная техника
6	Читальный зал	Персональные компьютеры