

ФЕДЕРАЛЬНОЕ АГЕНТСТВО СВЯЗИ

**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)

Первый проректор – проректор по учебной работе

Г.М. Машков
« 19 » 06 20 18 г.

Регистрационный №_18.07/32-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)

Анализ и проектирование бизнес-процессов 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», утвержденным приказом Министерства образования и науки Российской Федерации от 26.08.2016 № 1002, и в соответствии с рабочим учебным планом, утвержденным ректором университета.

1. Цели и задачи дисциплины

Целью преподавания дисциплины «Анализ данных» является:
сформировать у студентов системное представление о технологиях интеллектуального анализа данных, их применении и инструментах.

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

дать представление о моделях архитектуры хранилищ данных; познакомить с особенностями технологий 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.Б.12 является одной из дисциплин базовой части учебного плана подготовки бакалавриата по направлению «38.03.05 Бизнес-информатика». Исходный уровень знаний и умений, которыми должен обладать студент, приступая к изучению данной дисциплины, определяется изучением таких дисциплин, как «Информационное обеспечение бизнес-процессов предприятия»; «Информационные технологии в менеджменте»; «Прикладная статистика»

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

Процесс изучения дисциплины направлен на формирование следующих компетенций:

Компетенции, установленные ФГОС ВО

Таблица 1

№ п/п	Код компетенции	Наименование компетенции
1	ОПК-3	способностью работать с компьютером как средством управления информацией, работать с информацией из различных источников, в том числе в глобальных компьютерных сетях
2	ПК-17	способность использовать основные методы естественнонаучных дисциплин в профессиональной деятельности для теоретического и экспериментального исследования
3	ПК-18	способность использовать соответствующий математический аппарат и инструментальные средства для обработки, анализа и систематизации информации по теме исследования

Планируемые результаты обучения

Таблица 2

Код компетенции	знать	уметь	владеть
ОПК-3	основные методы, способы и средства получения, хранения, переработки информации	работать с компьютером как средством анализа информации, полученной из разных источников	технологиями создания и использования систем поддержки принятия решений на основе хранилищ данных

ПК-17	основные понятия математической статистики	использовать основные математические методы для теоретического и экспериментального исследования	методами количественного анализа и моделирования
ПК-18	методы и технологии Data Mining	использовать соответствующий математический аппарат и инструментальные средства для обработки, анализа и систематизации информации по теме исследования	технологиями интеллектуального анализа данных для решения организационно-экономических задач

4. Объем дисциплины и виды учебной работы

Очная форма обучения

Таблица 3

Вид учебной работы		Всего часов	Семестры
			5
Общая трудоемкость	5 ЗЕТ	180	180
Контактная работа с обучающимися		68.35	68.35
в том числе:			
Лекции		26	26
Практические занятия (ПЗ)		22	22
Лабораторные работы (ЛР)		18	18
Защита контрольной работы			-
Защита курсовой работы			-
Защита курсового проекта			-
Промежуточная аттестация		2.35	2.35
Самостоятельная работа обучающихся (СРС)		78	78
в том числе:			
Курсовая работа			-
Курсовой проект			-
И / или другие виды самостоятельной работы: подготовка к лабораторным работам, практическим занятиям, контрольным работам, изучение теоретического материала.		78	78
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. Системы оперативной обработки информации. Системы поддержки принятия решений	5		
2	Раздел 2. Хранилища данных	Концепция хранилища данных. Виды данных. Извлечение данных (ETL). Архитектуры хранилища данных. Управление жизненным циклом информации (ILM).	5		
3	Раздел 3. OLAP-системы	Технология OLAP (Online Analytical Processing). Архитектура OLAP-систем. Работа с измерениями	5		
4	Раздел 4. Интеллектуальный анализ данных	Задачи и применение Data Mining. Этапы процесса Data Mining.	5		
5	Раздел 5. Методы Data Mining	Классификация данных. Кластеризация данных. Прогнозирование данных. Поиск ассоциативных правил. Корреляционно-регрессионный анализ. Деревья решений. Искусственные нейронные сети. Карты Кохонена. Метод k-means. Алгоритм Apriori. Визуализация данных.	5		

5.2. Междисциплинарные связи с обеспечиваемыми (последующими) дисциплинами.

Таблица 5

№ п/п	Наименование обеспечиваемых (последующих) дисциплин
1	Маркетинг
2	Моделирование бизнес-процессов
3	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	4				14	18
2	Раздел 2. Хранилища данных	6	4			16	26
3	Раздел 3. OLAP-системы	4	2			16	22
4	Раздел 4. Интеллектуальный анализ данных	2		2		16	20
5	Раздел 5. Методы Data Mining	10	16	16		16	58
Итого:		26	22	18	-	78	144

6. Лабораторный практикум

Очная форма обучения

Таблица 7

№ п/п	Номер раздела (темы)	Наименование лабораторной работы	Всего часов
1	4	Знакомство с аналитической платформой Deductor. Предобработка данных	2
2	5	Классификация с помощью деревьев решений	2
3	5	Кластеризация данных	2
4	5	Корреляционный анализ. Выявление дубликатов и противоречий данных	2
5	5	Поиск ассоциативных правил	2
6	5	Предсказательная аналитика с помощью нейронной сети	4
7	5	Прогнозирование с помощью линейной регрессии	4
Итого:			18

7. Практические занятия (семинары)

Очная форма обучения

Таблица 8

№ п/п	Номер раздела (темы)	Наименование практических занятий (семинаров)	Всего часов
1	2	Анализ данных в многомерной модели	4
2	3	Работа с OLAP кубом	2
3	5	Анализ данных в реляционной модели	4
4	5	Индивидуальное задание. Анализ в криминалистике	2
5	5	Индивидуальное задание. Анализ розничных продаж	2
6	5	Индивидуальное задание. Анализ сельскохозяйственных предприятий	2
7	5	Индивидуальное задание. Анализ спортивных результатов	2
8	5	Индивидуальное задание. Анализ энергетической отрасли	2
9	5	Индивидуальное задание. Моделирование результатов деятельности агрофирмы	2
Итого:			22

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	14
2	Изучение теоретического материала. Подготовка к практическим и лабораторным занятиям	опрос	16

3	Изучение теоретического материала. Подготовка к практическим и лабораторным занятиям	опрос	16
4	Изучение теоретического материала. Подготовка к практическим и лабораторным занятиям	опрос	16
5	Изучение теоретического материала. Подготовка к практическим и лабораторным занятиям	опрос	16
Итого:			78

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 Методическими рекомендациями по формированию ФОС и приказом Минобрнауки России от 5 апреля 2017г. № 301 г. Москва "Об утверждении Порядка организации и осуществления образовательной деятельности по образовательным программам высшего образования - программам бакалавриата, программам специалитета, программам магистратуры" и является приложением 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.

- Фонд оценочных средств для проведения текущего контроля включают в себя тестовые задания (около 100 заданий).
- Фонд оценочных средств для проведения промежуточной аттестации включают в себя вопросы к зачету (около 30).

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, Федер. гос. образовательное бюджет. учреждение высш. проф. образования "С.-Петербург. гос. ун-т телекоммуникаций им. проф. М. А. Бонч-Бруевича". - СПб. : СПбГУТ, 2015. - 81 с. : ил. - Библиогр.: с. 81. - (в обл.) : 451.36 р.

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, Федеральное государственное образовательное бюджетное учреждение высшего профессионального образования "Санкт-Петербургский государственный университет телекоммуникаций им. проф. М. А. Бонч-Бруевича". - СПб. : СПбГУТ, 2013. - 47 с. : ил. -
2. Анализ данных. Часть 1. Подготовка данных к анализу [Электронный ресурс] : учебное пособие. - Москва : Московский городской педагогический университет, 2012. - . - ISBN 978-5-7786-0348-6. Анализ данных. Часть 1. Подготовка данных к анализу / Федин Ф. О. - 2012. - 204 с. - Б. ц. Книга находится в Премиум-версии ЭБС IPRbooks.
3. Анализ данных. Часть 2. Инструменты Data Mining [Электронный ресурс] : учебное пособие. - Москва : Московский городской педагогический университет, 2012. - . - ISBN 978-5-7786-0348-6. Анализ данных. Часть 2. Инструменты Data Mining / Федин Ф. О. - 2012. - 308 с. - Б. ц. Книга находится в Премиум-версии ЭБС IPRbooks.

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

14.2. Информационно-справочные системы:

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

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 к практическим занятиям, а также задания, вынесенные на самостоятельную работу, рекомендуется выполнять непосредственно после соответствующей темы лекционного курса, 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

Подготовку к каждому практическому занятию студент должен начать с ознакомления с планом практического занятия, который отражает содержание предложенной темы. Тщательное продумывание и изучение вопросов плана основывается на проработке текущего материала лекции, а затем изучения обязательной и дополнительной литературы, рекомендованной к данной теме. Все новые понятия по изучаемой теме необходимо выучить наизусть и внести в глоссарий, который целесообразно вести с самого начала изучения курса.

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.

В процессе подготовки к практическим занятиям, студентам необходимо обратить особое внимание на самостоятельное изучение рекомендованной

литературы. При всей полноте конспектирования лекции в ней невозможно изложить весь материал из-за лимита аудиторных часов. 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, опорными 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	Лаборатория	Лабораторное оборудование
4	Компьютерный класс	Персональные компьютеры
5	Аудитория для курсового и дипломного проектирования	Персональные компьютеры
6	Аудитория для самостоятельной работы	Персональные компьютеры
7	Читальный зал	Персональные компьютеры