

Аннотация к рабочей программе дисциплины  
**Б1.В.05 «Data Mining»**

**Объем трудоемкости:** 3 зачетных единицы.

**Цель дисциплины:** формирование у студентов знаний, умений и навыков (компетенций) по одному из приоритетных в современных информационных технологиях направлению – поиску и аналитической обработке данных.

**Задачи дисциплины:** формирование у бакалавров практических навыков сбора, систематизации и хранения данных; представления о технических и методологических средствах анализа данных, обеспечивающих сбор, хранение и систематизацию информации; ознакомление с основными принципами машинного обучения - а именно, видами задач машинного обучения, классами моделей (линейные, логические, нейросетевые), метриками качества и подходами к предварительной обработке данных.

**Место дисциплины в структуре образовательной программы**

Дисциплина Б1.В.05 «Data Mining» относится к части, формируемой участниками образовательных отношений Блока 1 «Дисциплины (модули)» учебного плана.

Перечень предшествующих дисциплин, необходимых для ее изучения:

- Информационно-коммуникационные технологии в профессиональной деятельности;
- Базы данных;
- Теория и технология программирования.

Перечень последующих дисциплин, для которых данная дисциплина является предшествующей в соответствии с учебным планом:

- Теория принятия решений;
- Технологическое предпринимательство;
- Анализ Big Data;
- Low-code аналитика;
- Аудит и оптимизация бизнес-процессов.

**Требования к уровню освоения дисциплины**

Изучение данной учебной дисциплины направлено на формирование у обучающихся следующих компетенций:

Код и наименование индикатора* достижения компетенции	Результаты обучения по дисциплине
<b>ПК-3 Способен регламентировать процессы подразделений организации и разрабатывать административные регламенты подразделений организации (в том числе кросс-функциональные процессы)</b>	
ИПК-3.1 Определяет источники, анализирует, собирает и систематизирует информацию для анализа	<i>Знает:</i> методы сбора информации (наблюдения, фиксация данных, хронометраж, фотография рабочего дня, техники проведения интервью и анкетирования, анализ документов и отчетной информации, изучение обратной связи от заинтересованных сторон) <i>Умеет:</i> выполнять наблюдения, интервью и анкетирование; анализировать, систематизировать и обобщать информацию; агрегировать, структурировать и обобщать информацию <i>Трудовое действие:</i> сбор информации о ходе и результатах процесса подразделения организации или административного регламента подразделения организации; оформление результатов сбора информации.
<b>ПК-2 Способен анализировать и исследовать большие данные с использованием существующей в организации методологической и технологической инфраструктуры</b>	

ИПК-2.2. Определяет источники, анализирует, собирает и систематизирует информацию для анализа	<p><i>Знает:</i> источники информации, в том числе информации, необходимой для обеспечения деятельности в предметной области заказчика исследования; виды источников данных: созданные человеком, созданные машинами; методы извлечения информации и знаний из гетерогенных, мультиструктурированных, неструктурированных источников, в том числе при потоковой обработке; режимы получения и обработки данных, поддержка режима реального времени.</p> <p><i>Умеет:</i> осуществлять взаимодействие с внутренними и внешними поставщиками данных из гетерогенных источников; использовать инструментальные средства для извлечения, преобразования, хранения и обработки данных из разнородных источников, в том числе в режиме реального времени; производить очистку данных для проведения аналитических работ.</p> <p><i>Трудовое действие:</i> извлечение, проверка и очистка больших объемов данных из гетерогенных источников; извлечение, проверка и очистка больших объемов данных из гетерогенных источников; оценка соответствия набора данных предметной области и задачам аналитических работ.</p>
<b>ПК-4 Способен обосновывать возможные решения и выбирать наиболее оптимальные</b>	
ИПК-4.2 Определяет источники, анализирует, собирает и систематизирует информацию для анализа	<p><i>Знает:</i> методы сбора, анализа, систематизации, хранения и поддержания в актуальном состоянии информации бизнес-анализа.</p> <p><i>Умеет:</i> проводить сбор и систематизацию информации; агрегировать, структурировать и обобщать; применять методы выбора оптимального решения.</p> <p><i>Трудовое действие:</i> анализ, обоснование и выбор решения.</p>

### Содержание дисциплины:

Распределение видов учебной работы и их трудоемкости по разделам дисциплины.  
Разделы (темы) дисциплины, изучаемые в 5 семестре (*очная форма обучения*).

№	Наименование разделов (тем)	Количество часов				
		Всего	Аудиторная работа			Внеаудиторная работа
			Л	ПЗ	ЛР	
1.	Данные: источники хранения данных в сети, современные подходы к обработке	8	2		2	4
2.	Методы сбора и систематизации информации	10	2		4	4
3.	Сбор данных из социальных сетей. API протоколы для VK, LinkedIn, Twitter, Мой Мир.	10	2		4	4
4.	API протокол для финансовых данных: Marketstack, Finnhub, IEX Cloud, Finam.	10	2		4	4
5.	Обработка данных. Библиотеки NumPy, Pandas.	12	2		4	6
6.	Визуализация информации. Библиотеки Matplotlib и SeaBorn. Онлайн-сервисы для построения дашбордов.	14	2		6	6
7.	Парсинг сайтов на Python.	10	2		2	6
8.	Вэб-аналитика: Google Analytics, Яндекс.Метрика.	10	2		2	6
9.	Модели машинного обучения на данных: модели регрессии, классификации, кластеризации.	17,8	2		6	9,8
	<i>ИТОГО по разделам дисциплины</i>	101, 8	18		34	49,8

	Контроль самостоятельной работы (КСР)	6				
	Промежуточная аттестация (ИКР)	0,2				
	Подготовка к текущему контролю					
	Общая трудоемкость по дисциплине	<b>108</b>				

**Курсовая работа:** не предусмотрена.

**Форма проведения аттестации по дисциплине:** зачет.

Автор: к.т.н., доцент Н.Ю. Нарыжная.