

## 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;
- Управление взаимодействиями в бизнес-среде;
- Статистические методы и модели.

#### 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-1 Способен руководить поставками инфокоммуникационных систем и (или) их составляющих</b>	
ИПК-1.3 Определяет источники, анализирует, собирает и систематизирует информацию для анализа	<i>Знает:</i> методы и инструменты поиска информации; способы поиска информации по продажам инфокоммуникационных систем и (или) их составляющих <i>Умеет:</i> 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; осуществлять долгосрочное прогнозирование продаж инфокоммуникационных систем и (или) их составляющих; проводить анализ и распространение лучших практик продаж инфокоммуникационных систем и (или) их составляющих <i>Трудовое действие:</i> извлечение, проверка и очистка данных из гетерогенных источников; поиск и обработка информации по продажам инфокоммуникационных систем и (или) их составляющих.
<b>ПК-2 Способен разрабатывать отдельные функциональные направления управления рисками</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 описания и актуализации рисков; методы сбора и систематизации этой информации</p> <p><i>Умеет:</i> использовать инструментальные средства и методы для сбора и обработки информации по актуализации рисков, оценке качества их описания, в том числе в режиме реального времени; составлять аналитические отчеты с информацией о рисках для руководителя подразделения и ответственных за мероприятия по рискам работников</p> <p><i>Трудовое действие:</i> сбор информации, контроль качества работы сотрудников по описанию и актуализации рисков; обеспечение информацией текущего управления рисками на постоянной основе; представление аналитической информации о рисках руководителю подразделения и ответственным за мероприятия по рискам работникам</p>
<b>ПК-4 Способен обосновывать возможные решения и выбирать наиболее оптимальные</b>	
ИПК-4.1 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; информационные технологии (программное обеспечение), применяемые для целей сбора и систематизации информации;</p> <p><i>Умеет:</i> проводить сбор и систематизацию информации; агрегировать, структурировать и обобщать данные; готовить обоснование для выбора решения</p> <p><i>Трудовое действие:</i> определение источников данных для анализа, идентификация внешних и внутренних источников данных для проведения аналитических работ по обоснованию и выбору наиболее оптимального 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.