

**Учреждение образования
«Гродненский государственный аграрный университет»**

УТВЕРЖДАЮ

Ректор учреждения образования
«Гродненский государственный
аграрный университет»

_____ В.К. Пестис

«__» _____ 2019 г.

Регистрационный № УД-_____/уч.

ТЕХНОЛОГИИ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ДАННЫХ

Учебная программа учреждения высшего образования для магистрантов

1-25 80 01 «Экономика»(профилизация – «Аграрная экономика»)

2019 г.

СОСТАВИТЕЛИ:

Т. Н. Изосимова, заведующий кафедрой информатики и экономико-математического моделирования в агропромышленном комплексе Учреждения образования «Гродненский государственный аграрный университет», кандидат физико-математических наук, доцент

РЕЦЕНЗЕНТЫ:

Л.В.Рудикова, заведующий кафедрой современных технологий программирования Учреждения образования «Гродненский государственный университет им. Я.Купалы», кандидат физико-математических наук, доцент;

Дорошкевич И.Н., проректор по учебной работе Учреждения образования «Гродненский государственный аграрный университет», кандидат экономических наук, доцент

РЕКОМЕНДОВАНА К УТВЕРЖДЕНИЮ:

Кафедрой информатики и экономико-математического моделирования в агропромышленном комплексе

(протокол № 10 от 24.05.2019);

Методическим советом Учреждения образования «Гродненский государственный аграрный университет»

(протокол № от)

Ответственный за редакцию Т.Н.Изосимова

Ответственный за выпуск Т.В.Снопко

ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

Цели и задачи учебной дисциплины

Цель изучения дисциплины – формирование у магистрантов знаний и навыков по использованию технологий высокопроизводительной интеллектуальной аналитической обработки информации, направленной на то, чтобы оперативно извлекать из значительных массивов накопленных и поступающих данных ценные экспертные знания, поддерживая эффективную управленческую деятельность.

Задачи дисциплины:

- дать основные теоретические знания о базовых понятиях и методах интеллектуального анализа данных;
- выработать навыки и умения по использованию программных средств на основе технологии интеллектуального анализа данных для решения практических задач.

Место учебной дисциплины в системе подготовки специалиста с высшим образованием

Дисциплина «Технологии интеллектуального анализа данных» включена в «Государственный компонент» учебного плана для магистрантов по специальности 1-25 80 01 «Экономика» (профилизация «Аграрная экономика»).

Освоение дисциплины базируется на компетенциях, приобретенных ранее магистрантами при изучении дисциплин «Высшая математика», «Компьютерные информационные технологии», «Статистика».

Требования к освоению учебной дисциплины

В результате изучения дисциплины магистрант должен закрепить и развить следующие академические (АК) и социально-личностные (СЛК) компетенции:

АК-1. Уметь применять базовые научно-теоретические знания для решения теоретических и практических профессиональных задач.

АК-2. Владеть сравнительным анализом.

АК-3. Уметь работать самостоятельно.

АК-4. Владеть междисциплинарным подходом для решения задач.

АК-5. Иметь навыки при работе с компьютерной техникой.

СЛК-1. Обладать способностью к межличностным коммуникациям.

СЛК-2. Уметь работать в коллективе.

В результате изучения дисциплины магистрант должен обладать следующими профессиональными компетенциями (ПК):

ПК-1. Принимать решения о выборе оптимальной информационной структуры производства.

ПК-2. Проводить организационно-управленческое обследование производства.

ПК-3. Выбирать наиболее эффективные формы адаптации предприятия к условиям внешней среды.

ПК-4. Использовать современные информационные технологии и прикладные программы обработки данных для обоснования управленческих решений.

ПК-5. Принимать эффективные управленческие решения в экономической сфере.

ПК-6. Уметь анализировать деятельность организации на международных рынках и разрабатывать стратегии их завоевания.

ПК-7. Оценивать результаты хозяйственной деятельности организаций.

ПК-8. Работать с научной, технической и патентной литературой.

Для приобретения профессиональных компетенций ПК-1 – ПК-8 в результате изучения дисциплины магистрант должен знать:

- задачи систем поддержки принятия решений;
- задачи Data Mining;
- практическое применение Data Mining;
- модели Data Mining;
- методы Data Mining;
- процесс обнаружения знаний;
- средства Data Mining;
- сферы применения Data Mining.

Уметь и быть способным:

- квалифицировать задачи Data Mining;
- применять средства интеллектуального анализа данных.

Общее количество часов и количество аудиторных часов

На усвоение дисциплины максимально отводится 100 часов, в том числе 42 часа аудиторных.

Форма получения высшего образования

При изучении дисциплины предусмотрены две формы обучения: дневная и заочная.

Распределение аудиторного времени по видам занятий, курсам и семестрам

Дисциплина изучается в 1 семестре, при этом аудиторное время делится на 18 часов лекционных и 24 часа лабораторных занятий.

Формы текущей аттестации по учебной дисциплине

Оценка учебных достижений магистранта проводится на зачете. При этом используются критерии, утвержденные Министерством образования Республики Беларусь.

Оценка промежуточных учебных достижений магистранта осуществляется в соответствии с избранной кафедрой шкалой оценок.

Для оценки достижений магистранта используется следующий диагностический инструментарий (проверяются компетенции):

- выступление магистрантов на конференции по подготовленному реферату (АК-1, АК-3, АК-4, АК-5, СЛК-1, СЛК-2, ПК-6);
- проведение текущих контрольных опросов по отдельным темам (ПК-1 – ПК-7);
- защита выполненных на лабораторно-практических занятиях индивидуальных заданий (АК-3, АК-5, СЛК-1, ПК-1– ПК-7);
- сдача зачета по дисциплине (АК-1 – АК-5, ПК-1 – ПК-7).

СОДЕРЖАНИЕ УЧЕБНОГО МАТЕРИАЛА

Тема 1. Системы поддержки принятия решений

Задачи систем поддержки принятия решений. Базы данных основа СППР. Хранилища данных. OLAP-системы

Тема 2. Основные понятия интеллектуального анализа данных

Задачи Data Mining. Практическое применение Data Mining. Модели Data Mining. Методы Data Mining. Процесс обнаружения знаний. Средства Data Mining.

Тема 3. Основные методы интеллектуального анализа данных

3.1 Классификация и регрессия.

Постановка задачи. Представление результатов: *Правила классификации. Деревья решений. Математические функции.*

Методы построения правил классификации *Алгоритм построения 1-правил. Метод Naïve Bayes.*

Методы построения деревьев решений: *Методика "разделяй и властвуй". Алгоритм покрытия.*

Методы построения математических функций *Общий вид. Линейные методы. Метод наименьших квадратов. Нелинейные методы Support Vector Machines (SVM). Регуляризационные сети (Regularization Networks.) Дискретизации и редкие сетки.*

Прогнозирование временных рядов. *Постановка задач. Методы прогнозирования временных рядов.*

3.2 Поиск ассоциативных правил

Постановка задачи. *Формальная постановка задачи. Секвенциальный анализ. Разновидности задачи поиска ассоциативных правил.*

Представление результатов.

Алгоритмы. *Алгоритм Apriori. Разновидности алгоритма Apriori.*

3.3 Кластеризация

Постановка задачи кластеризации. Формальная постановка задачи. Меры близости, основанные на расстояниях, используемые в алгоритмах кластеризации.

Представление результатов.

Базовые алгоритмы кластеризации. *Классификация алгоритмов. Иерархические алгоритмы. Неиерархические алгоритмы.*

Адаптивные методы кластеризации. *Выбор наилучшего решения и качество кластеризации. Использование формальных критериев качества в адаптивной кластеризации. Пример адаптивной кластеризации.*

3.4 Нейронные сети.

Область применения. Элементы нейронных сетей. Архитектура нейронных сетей.

Обучение нейронных сетей. Модели нейронных сетей. Карты Кохонена. Программное обеспечение для работы с нейронными сетями.

3.5 Визуальный анализ данных – Visual Mining

Выполнение визуального анализа данных. Характеристики средств визуализации данных. Методы визуализации.

Тема 4. Инструментальные средства интеллектуального анализа данных

Инструменты Data Mining и их классификация. Использование компьютерных технологий для решения задач Data Mining.

УЧЕБНО-МЕТОДИЧЕСКАЯ КАРТА

Номер раздела, темы, занятия	Название раздела, темы, занятия; перечень изучаемых вопросов	Всего часов	Количество аудиторных часов				Количество часов, выделяемых на самостоятельную работу студентов (в т.ч. часы, выделяемые на выполнение курсовой работы/проект)	Материальное обеспечение занятия (наглядные, методические пособия и др.)	Литература	Формы контроля знаний
			лекции	практические (семинарские) занятия	лабораторные занятия	самостоятельная работа студента (КСР)				
1	2	3	4	5	6	7	8	9	10	11
<i>Тема 1.</i>	Системы поддержки принятия решений Задачи систем поддержки принятия решений. Базы данных основа СППР. Хранилища данных. OLAP-системы	10	2			8		ЭУМК и электронные учебники, размещенные в локальной сети библиотеки, тестирование через сайт дистанционного обучения	[1]	Тестирование
<i>Тема 2.</i>	Основные понятия интеллектуального анализа данных Задачи Data Mining. Практическое применение Data Mining. Модели Data Mining. Методы Data Mining. Процесс обнаружения знаний. Средства Data Mining	12	2			10		ЭУМК и электронные учебники, размещенные в локальной сети библиотеки, тестирование через сайт дистанционного обучения	[1]	Тестирование
<i>Тема 3.</i>	Основные методы интеллектуального анализа данных	44	14			30		ЭУМК и электронные учебники, размещенные в локальной сети библиотеки, тестирование через сайт дистанционного обучения		Тестирование
<i>3.1</i>	Классификация и регрессия		4			6			[1, 8]	

1	2	3	4	5	6	7	8	9	10	11
3.1.1	Постановка задачи. Представление результатов: <i>Правила классификации. Деревья решений. Математические функции.</i> Методы построения правил классификации. <i>Алгоритм построения 1-правил</i> <i>Метод Naive Bayes.</i> Методы построения деревьев решений: <i>Методика "разделяй и властвуй". Алгоритм покрытия.</i>		2							
3.1.2	Методы построения математических функций <i>Общий вид</i> <i>Линейные методы. Метод наименьших квадратов. Нелинейные методы Support Vector Machines (SVM). Регуляризационные сети (Regularization Networks). Дискретизации и редкие сетки.</i> Прогнозирование временных рядов. <i>Постановка задач. Методы прогнозирования временных рядов.</i>		2							
3.2.	Поиск ассоциативных правил Постановка задачи. <i>Формальная постановка задачи. Секвенциальный анализ. Разновидности задачи поиска ассоциативных правил.</i> Представление результатов. Алгоритмы. <i>Алгоритм Apriori. Разновидности алгоритма Apriori</i>		2			6				
3.3	Кластеризация		4			6			[1, 2]	
3.3.1	Постановка задачи кластеризации. <i>Формальная постановка задачи. Меры близости, основанные на расстояниях, используемые в алгоритмах кластеризации.</i> Представление результатов. Базовые алгоритмы кластеризации. <i>Классификация алгоритмов. Иерархические алгоритмы. Неиерархические алгоритмы.</i>		2							
3.3.2	Адаптивные методы кластеризации. <i>Выбор наилучшего решения и качество кластеризации. Использование формальных критериев качества в адаптивной кластеризации. Пример адаптивной кластеризации.</i>		2							
3.4.	Нейронные сети Область применения. <i>Элементы нейронных сетей. Архитектура нейронных сетей. Обучение нейронных сетей. Модели нейронных сетей. Программное обеспечение для работы с нейронными сетями.</i>		2			6			[1, 3]	

1	2	3	4	5	6	7	8	9	10	11
3.5	Визуальный анализ данных – Visual Mining Выполнение визуального анализа данных. Характеристики средств визуализации данных. Методы визуализации.		2			6			[1]	
Тема 4	Инструментальные средства интеллектуального анализа данных Инструменты Data Mining и их классификация. Использование компьютерных технологий для решения задач Data Mining.	34			24	10		ЭУМК и электронные учебники, размещенные в локальной сети библиотеки, тестирование через сайт дистанционного обучения	[1,2,3,4-8]	Тестирование
4.1	Классификация и регрессия				8	2				
4.1.1	<i>Алгоритмы метода деревьев решений</i>				2					Отчёт по лаб.работе
4.1.2	<i>Метод опорных векторов</i>				2					Отчёт по лаб.работе
4.1.3	<i>Метод "ближайшего соседа"</i>				2					Отчёт по лаб.работе
4.1.4	<i>Байесовская классификация</i>				2					Отчёт по лаб.работе
4.2	Оценивания и прогнозирование				4	2				
4.2.1	<i>Прогнозирование и временные ряды</i>				2					Отчёт по лаб.работе
4.2.2	<i>Прогнозирование и визуализация</i>				2					Отчёт по лаб.работе
4.3	Поиска ассоциативных правил				4	2				Отчёт по лаб.работе
4.4	Кластеризация и сегментация				4	2				
4.4.1	<i>Иерархические методы</i>				2					Отчёт по лаб.работе
4.4.2	<i>Итеративные методы</i>				2					Отчёт по лаб.работе
4.5	Нейронные сети				4	2				Отчёт по лаб.работе
	Итого	100	18		24	58				

УЧЕБНО-МЕТОДИЧЕСКАЯ КАРТА (заочное)

Номер раздела, темы, занятия	Название раздела, темы, занятия; перечень изучаемых вопросов	Всего часов	Количество аудиторных часов				Количество часов, выделяемых на самостоятельную работу студентов (в т.ч. часы, выделяемые на выполнение курсовой работы/проект)	Материальное обеспечение занятия (наглядные, методические пособия и др.)	Литература	Формы контроля знаний
			лекции	практические (семинарские) занятия	лабораторные занятия	самостоятельная работа студента (КСР)				
1	2	3	4	5	6	7	8	9	10	11
<i>Тема 1.</i>	Системы поддержки принятия решений Задачи систем поддержки принятия решений. Базы данных основа СППР. Хранилища данных. OLAP-системы	10				10		ЭУМК и электронные учебники, размещенные в локальной сети библиотеки, тестирование через сайт дистанционного обучения	[1]	Тестирование
<i>Тема 2.</i>	Основные понятия интеллектуального анализа данных Задачи Data Mining. Практическое применение Data Mining. Модели Data Mining. Методы Data Mining. Процесс обнаружения знаний. Средства Data Mining	10				10		ЭУМК и электронные учебники, размещенные в локальной сети библиотеки, тестирование через сайт дистанционного обучения	[1]	Тестирование
<i>Тема 3.</i>	Основные методы интеллектуального анализа данных	44	4			40		ЭУМК и электронные учебники, размещенные в локальной сети библиотеки, тестирование через сайт дистанционного обучения		Тестирование

1	2	3	4	5	6	7	8	9	10	11
3.1	Классификация и регрессия		1			10			[1, 8]	
3.1.1	Постановка задачи. Представление результатов: <i>Правила классификации. Деревья решений. Математические функции.</i> Методы построения правил классификации. <i>Алгоритм построения 1-правил</i> <i>Метод Naive Bayes.</i> Методы построения деревьев решений: <i>Методика "разделяй и властвуй". Алгоритм покрытия.</i>									
3.1.2	Методы построения математических функций <i>Общий вид</i> <i>Линейные методы. Метод наименьших квадратов. Нелинейные методы Support Vector Machines (SVM). Регуляризационные сети (Regularization Networks). Дискретизации и редкие сетки.</i> Прогнозирование временных рядов. <i>Постановка задач. Методы прогнозирования временных рядов.</i>									
3.2.	Поиск ассоциативных правил Постановка задачи. <i>Формальная постановка задачи. Секвенциальный анализ. Разновидности задачи поиска ассоциативных правил.</i> Представление результатов. Алгоритмы. <i>Алгоритм Apriori. Разновидности алгоритма Apriori</i>									
3.3	Кластеризация		1			10			[1, 2]	
3.3.1	Постановка задачи кластеризации. <i>Формальная постановка задачи. Меры близости, основанные на расстояниях, используемые в алгоритмах кластеризации.</i> Представление результатов. Базовые алгоритмы кластеризации. <i>Классификация алгоритмов. Иерархические алгоритмы. Неиерархические алгоритмы</i>									
3.3.2	Адаптивные методы кластеризации. <i>Выбор наилучшего решения и качество кластеризации. Использование формальных критериев качества в адаптивной кластеризации. Пример адаптивной кластеризации.</i>		2							

1	2	3	4	5	6	7	8	9	10	11
3.4.	Нейронные сети Область применения. Элементы нейронных сетей. Архитектура нейронных сетей. Обучение нейронных сетей. Модели нейронных сетей. Программное обеспечение для работы с нейронными сетями.		1			10			[1, 3]	
3.5	Визуальный анализ данных – Visual Mining Выполнение визуального анализа данных. Характеристики средств визуализации данных. Методы визуализации.		1			10			[1]	
<i>Тема 4</i>	Инструментальные средства интеллектуального анализа данных Инструменты Data Mining и их классификация. Использование компьютерных технологий для решения задач Data Mining.	36			6	30		ЭУМК и электронные учебники, размещенные в локальной сети библиотеки, тестирование через сайт дистанционного обучения	[1,2,3,4-8]	Тестирование
4.1	Классификация и регрессия				1					Отчёт по лаб.работе
4.1.1	<i>Алгоритмы метода деревьев решений</i>									
4.1.2	<i>Метод опорных векторов</i>									
4.1.3	<i>Метод "ближайшего соседа"</i>									
4.1.4	<i>Байесовская классификация</i>									
4.2	Оценивания и прогнозирование				1					Отчёт по лаб.работе
4.2.1	<i>Прогнозирование и временные ряды</i>									
4.2.2	<i>Прогнозирование и визуализация</i>									
4.3	Поиска ассоциативных правил				1					Отчёт по лаб.работе
4.4	Кластеризация и сегментация				1					Отчёт по лаб.работе
4.4.1	<i>Иерархические методы</i>									
4.4.2	<i>Итеративные методы</i>									
4.5	Нейронные сети				2					Отчёт по лаб.работе
	Итого	100	4		6	90				

ИНФОРМАЦИОННО-МЕТОДИЧЕСКАЯ ЧАСТЬ

Контроль учебных достижений магистрантов

Для текущего контроля учебных достижений магистрантов используются тесты, разноуровневые контрольные задания, устный опрос во время занятий и другие средства диагностики. Итоговая оценка учебных достижений проводится на зачете. Для этого используются критерии, утвержденные Министерством образования Республики Беларусь.

Примерный перечень вопросов для итогового контроля

1. Расскажите о задачах систем поддержки принятия решений.
2. База данных как основа систем поддержки принятия решений.
3. Расскажите об архитектуре систем поддержки принятия решений.
4. Расскажите о концепции хранилища данных.
5. Расскажите об организации хранилища данных.
6. Расскажите о многомерной модели данных.
7. Дайте определение OLAP-системе, основные понятия.
8. Расскажите о концептуальном многомерном представлении данных.
9. Расскажите об архитектуре OLAP-систем.
10. Расскажите об основных понятиях анализа данных.
11. Расскажите о задачах анализа данных и областях практического применения.
12. Перечислите основные этапы построения моделей.
13. Расскажите о методике анализа данных.
14. Перечислите инструментальные средства для интеллектуального анализа баз данных.
15. Расскажите о выявлении закономерностей и прогнозировании нейросетевыми и другими современными методами.
16. Расскажите о задаче классификации и регрессии. Сформулируйте ее постановку, расскажите об алгоритмах решения, представлении результатов.
17. Сформулируйте задачу поиска ассоциативных правил и её постановку. Расскажите об алгоритмах решения, представлении результатов.
18. Расскажите о задаче кластеризации. Сформулируйте постановку задачи, расскажите об алгоритмах решения, представлении результатов.
19. Расскажите о выполнении визуального анализа данных и методах визуализации

Перечень рекомендуемой литературы

Основная литература:

1. Барсегян, А. А. Анализ данных и процессов: учеб. пособие / А. А. Барсегян, М. С. Куприянов, И. И. Холод, М. Д. Тесс, С. И. Елизаров. – 3-е изд., перераб. и доп. – СПб.: БХВ-Петербург, 2009. – 512 с.:
2. Мандель, И.Д. Кластерный анализ / И.Д. Мандель –М., 1988. – 176 с.
3. Медведев, В. С. Нейронные сети. MATLAB 6 / В. С. Медведев, В. Г. Потемкин ; под общ. ред. В. Г. Потемкина. – М., 2002. – 496 с.

Дополнительная литература:

4. Дюк, В. Data Mining: учебный курс / В. Дюк, А. Самойленко. –СПб., 2001. 368 с.

5. Дронов, С. В. Многомерный статистический анализ : учеб. пособие / С. В. Дронов. – Барнаул, 2003. – 213 с.
6. Лагутин, М. Б. Наглядная математическая статистика : учеб. пособие / М. Б. Лагутин. – М., 2007. – 472 с.
7. Информационные аналитические системы [Электронный ресурс] : учебник / Т. В. Алексеева, Ю. В. Амириди, В. В. Дик и др.; под ред. В. В. Дика. - М.: МФПУ Синергия, 2013. - 384 с. - Режим доступа по паролю: <http://znanium.com/bookread2.php?book=451186> (ЭБС Znanium)
8. Интеллектуальный анализ временных рядов: Учебное пособие [Электронный ресурс] / Н.Г. Ярушкина, Т.В. Афанасьева, И.Г. Перфильева. - М.: ИД ФОРУМ: ИНФРА-М, 2012. - Режим доступа по паролю: <http://znanium.com/bookread2.php?book=249314> (ЭБС Znanium)

ПРОТОКОЛ СОГЛАСОВАНИЯ УЧЕБНОЙ ПРОГРАММЫ

Название учебной дисциплины, с которой требуется согласование	Название кафедры	Предложения об изменениях в содержании учебной программы по изучаемой учебной дисциплине	Решение, принятое кафедрой, разработавшей учебную программу (с указанием даты и номера протокола)
Количественные методы анализа АПК	Кафедра информатики и ЭММ в АПК	Нет предложений	Рассмотрено на заседании кафедры, протокол №10 от 24 мая 2019 г.
Моделирование и оптимизация бизнес-процессов в АПК	Кафедра информатики и ЭММ в АПК	Нет предложений	Рассмотрено на заседании кафедры, протокол №10 от 24 мая 2019 г.

ДОПОЛНЕНИЯ И ИЗМЕНЕНИЯ К УЧЕБНОЙ ПРОГРАММЕ УВО на ____/____ учебный год

№ п/п	Дополнения и изменения	Основание

Учебная программа пересмотрена и одобрена на заседании кафедры информатики и ЭММ в АПК

_____ (название кафедры) (протокол № ____ от _____ 201_ г.)

Заведующий кафедрой
кандидат физ.-мат. наук, доцент _____ Т. Н. Изосимова
(ученая степень, ученое звание) (подпись) (И.О.Фамилия)

УТВЕРЖДАЮ
Декан факультета
кандидат эк. наук, доцент _____ А. В. Грибов
(ученая степень, ученое звание) (подпись) (И.О.Фамилия)ⁱ