

Белорусский государственный университет



Проректор по учебной работе и образовательным инновациям

О. И. Чуприс

2018 г.

Регистрационный № УД- 6240 /уч.

РАСПОЗНАВАНИЕ ОБРАЗОВ

Учебная программа учреждения высшего образования по учебной дисциплине для специальности первой ступени высшего образования:

**1-31 03 03 Прикладная математика (по направлениям)
направления специальности**

1-31 03 03 - 01 Прикладная математика (научно-производственная деятельность)

2018 г.

Учебная программа составлена на основе образовательного стандарта высшего образования ОСВО 1-31 03 03-2013 и учебного плана G31-173/уч. от 30.05.2013, G31и-190/уч. от 30.05.2013

СОСТАВИТЕЛЬ:

Е.Е. Жук, профессор кафедры математического моделирования и анализа данных факультета прикладной математики и информатики Белорусского государственного университета, доктор физико-математических наук, профессор.

РЕКОМЕНДОВАНА К УТВЕРЖДЕНИЮ:

Кафедрой математического моделирования и анализа данных Белорусского государственного университета (протокол № 13 от 29 марта 2018);

Научно-методическим Советом Белорусского государственного университета (протокол № 5 от 4 мая 2018 г.).



 /Богачен Н.А., зав. кафедрой ММАФ/

ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

Распознавание образов – это раздел математики, который занимается изучением методов статистической классификации многомерных наблюдений для различных прикладных областей, где требуется принятие решений: в технических исследованиях, экономике, теории и практике управления, социологии, медицине и т.д. С ними имеют дело специалисты во многих отраслях практической деятельности.

В основу положен принцип описания наблюдений из различных классов при помощи своих вероятностных распределений. При этом используются вероятностно-статистический математический аппарат.

Современное развитие вычислительной техники и программного обеспечения благоприятствует широкому внедрению методов статистической классификации в практику. Пакеты прикладных программ (ППП) с удобным пользовательским интерфейсом (такие как ППП STATISTICA), а также современные технологии компьютерного анализа данных (например, система R), способствуют их применению.

Учебная дисциплина «Распознавание образов» знакомит студентов с классическими и современными методами решения задач статистической классификации.

Целью преподавания учебной дисциплины является:

- изучение основных математических моделей, статистических методов, алгоритмов и программного обеспечения, используемых для принятия решений;
- расширение фундаментального математического образования;
- формирование у студентов профессиональных знаний в предметной области, подкрепленных владением математическими методами и пакетами прикладных программ.

Задачи учебной дисциплины состоят в следующем:

- изучение основных понятий теории статистической классификации (риск, байесовское решающее правило, модель Фишера и многое другое);
- изучение методов и алгоритмов статистической классификации при наличии обучения (дискриминантный анализ);
- изучение методов и алгоритмов статистического анализа неклассифицированных выборок (кластер-анализ).

Учебная дисциплина «Распознавание образов» относится к циклу дисциплин специализаций.

Учебная программа составлена с учетом межпредметных связей с учебными дисциплинами. Так, основой для изучения дисциплины «Распознавание образов» является дисциплина «Теория вероятностей и математическая статистика». Знания, полученные в результате изучения дисциплины, будут использованы при изучении дисциплин «Имитационное и

статистическое моделирование», «Модели и методы теории прогнозирования».

В результате освоения учебной дисциплины студент должен:

знать:

- методы дискриминантного анализа;
- методы кластер-анализа;
- методы формирования информативных признаков;

уметь:

- осуществлять дискриминантный анализ и классифицировать вновь поступившие наблюдения;
- проводить кластерный анализ и определять при этом неизвестное число классов на основе различных методов (например, иерархической классификации);
- осуществлять сжатие данных и выделять информативные признаки;
- использовать соответствующие пакеты прикладных программ, а также современные технологии компьютерного анализа данных;
- оценивать эффективность принимаемых статистических решений;

владеть:

- методами решения основных задач статистической классификации;
- навыками по сбору и подготовке статистических данных, подлежащих классификации;
- навыками применения современных ППП для решения задач статистической классификации многомерных данных;
- навыками по подготовке отчетов с результатами проведенного статистического анализа, интерпретацией полученных результатов, а также комментариями, выводами и рекомендациями.

Освоение учебной дисциплины «Распознавание образов» должно обеспечить формирование следующих академических, социально-личностных и профессиональных компетенций:

академические компетенции:

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

АК-2. Владеть системным и сравнительным анализом.

АК-3. Владеть исследовательскими навыками.

АК-4. Уметь работать самостоятельно.

АК-5. Быть способным вырабатывать новые идеи (обладать креативностью).

АК-6. Владеть междисциплинарным подходом при решении проблем.

АК-7. Иметь навыки, связанные с использованием технических устройств, управлением информацией и работой с компьютером.

АК-9. Уметь учиться, повышать свою квалификацию в течение всей жизни.

социально-личностные компетенции:

СЛК-2. Быть способным к социальному взаимодействию.

СЛК-3. Обладать способностью к межличностным коммуникациям.

СЛК-6. Уметь работать в команде.

профессиональные компетенции:

ПК-19. Взаимодействовать со специалистами смежных профилей.

ПК-23. Владеть современными средствами телекоммуникаций.

Структура содержания учебной дисциплины включает такие дидактические единицы, как темы (разделы), в соответствии с которыми разрабатываются и реализуются соответствующие лекционные и лабораторные занятия. Примерная тематика занятий приведена в информационно-методической части.

Дисциплина изучается в 7 семестре. Всего на освоение учебной дисциплины «Распознавание образов» отведено всего 159 часов, из них 68 аудиторных часов, в том числе лекций – 34 часа, лабораторных занятий – 30 часов, УСП – 4 часа.

Трудоемкость учебной дисциплины составляет 4 зачетные единицы.

Форма текущей аттестации – зачет, экзамен.

СОДЕРЖАНИЕ УЧЕБНОГО МАТЕРИАЛА

Раздел 1. Введение

Тема 1.1. Предмет курса. Основные понятия прикладной теории статистических решений. Вероятностная математическая модель и задача статистической классификации. Дискриминантный и кластерный анализ.

Раздел 2. Оптимальная статистическая классификация при известных характеристиках классов

Тема 2.1. Оптимальное (байесовское) решающее правило и его риск. Риск как мера оптимальности решающих правил и его свойства. Оптимальное (байесовское) решающее правило и его риск.

Тема 2.2. Решающие правила в случае двух классов. Байесовское решающее правило в случае двух классов и его риск. Решающее правило Зигерта-Котельникова.

Тема 2.3. Статистическая классификация многомерных нормальных наблюдений. Статистическая классификация многомерных гауссовских (нормальных) наблюдений. Байесовское решающее правило и его риск для модели Фишера в случае двух классов.

Раздел 3. Дискриминантный анализ

Тема 3.1. Подстановочные решающие правила в дискриминантном анализе и их риск. Подстановочные решающие правила в дискриминантном анализе. Подстановочный риск и его строгая состоятельность. Подстановочное байесовское решающее правило в дискриминантном анализе и свойства его риска.

Тема 3.2. Непараметрический дискриминантный анализ. Непараметрический дискриминантный анализ на основе различных оценок плотности распределения вероятностей. Непараметрический дискриминантный анализ. Решающее правило по методу k -ближайших соседей (k -NN-классификатор).

Тема 3.3. Параметрический дискриминантный анализ. Параметрический дискриминантный анализ на основе подстановочного байесовского решающего правила. Дискриминантный анализ многомерных гауссовских наблюдений.

Раздел 4. Кластер-анализ

Тема 4.1. Вероятностная постановка задачи кластер-анализа. Вероятностная постановка задачи кластер-анализа и условия ее разрешимости.

Тема 4.2. Прямые методы кластер-анализа. Метод моментов и метод минимума χ^2 -статистики. Метод максимального правдоподобия и *ЕМ*-алгоритм.

Тема 4.3. Методы кластер-анализа с обратной связью. Метод кластер-анализа с обратной связью, основанный на обобщенной функции правдоподобия. Метод кластер-анализа с обратной связью, основанный на гипотезе компактности (метод *L*-средних).

Тема 4.4. Кластер-анализ при неизвестном числе классов. Различные методы кластер-анализа в условиях неизвестного числа классов. Иерархический кластер-анализ.

Раздел 5. Формирование информативных признаков в задачах статистической классификации

Тема 5.1. Формирование информативных признаков на основе риска. Сжатие наблюдений на основе критерия минимума риска.

Тема 5.2. Формирование информативных признаков методом главных компонент. Сжатие данных методом главных компонент. Свойства и интерпретация главных компонент в модели статистической классификации.

УЧЕБНО-МЕТОДИЧЕСКАЯ КАРТА УЧЕБНОЙ ДИСЦИПЛИНЫ

Но мер раз дел а, тем ы	Название раздела, темы	Количество аудиторных часов			Коли честв о часов УСР	Форма контроля знаний
		Лек ции	Семи нарск ие занят ия	Лабо ратор ные занят ия		
1	Введение	2		2		
	Предмет курса	2		2		Устный опрос
2	Оптимальная статистическая классификация при известных характеристиках классов	6		6		
2.1	Оптимальное (байесовское) решающее правило и его риск	2		2		Устный опрос
2.2	Решающие правила в случае двух классов	2		2		Защита подготовленного студентом реферата
2.3	Статистическая классификация многомерных нормальных наблюдений	2		2		Отчет по заданию с устной защитой
3	Дискриминантный анализ	10		8	2	
3.1	Подстановочные решающие правила в дискриминантном анализе и их риск	2		2		Устный опрос. Отчет по заданию с устной защитой
3.2	Непараметрический дискриминантный анализ	4		2	2	Отчет по заданию с устной защитой
3.3	Параметрический дискриминантный анализ	4		4		Коллоквиум
4	Кластер-анализ	12		10	2	
4.1	Вероятностная постановка задачи кластер-анализа	2		2		Защита подготовленного студентом реферата

4.2	Прямые методы кластер-анализа	4		2	2	Отчет по заданию с устной защитой
4.3	Методы кластер-анализа с обратной связью	4		4		Устный опрос
4.4	Кластер-анализ при неизвестном числе классов	2		2		Отчет по заданию с устной защитой
5	Формирование информативных признаков в задачах статистической классификации	4		4		
5.1	Формирование информативных признаков на основе риска	2		2		Отчет по заданию с устной защитой
5.2	Формирование информативных признаков методом главных компонент	2		2		Коллоквиум
ИТОГО		34		30	4	

ИНФОРМАЦИОННО-МЕТОДИЧЕСКАЯ ЧАСТЬ

Перечень основной литературы

1. Андерсон, Т. В. Введение в многомерный статистический анализ / Т. В. Андерсон. – М. : Физматгиз, 1963.
2. Прикладная статистика. Классификация и снижение размерности / С. А. Айвазян [и др.]. – М., Финансы и статистика, 1989.
3. Фукунага, К. Введение в статистическую теорию распознавания образов / К. Фукунага. – М. : Наука, 1979 .
4. Харин, Ю. С., Математическая и прикладная статистика / Ю. С. Харин, Е. Е. Жук. – Мн. : БГУ, 2005.
5. Харин, Ю. С. Теория вероятностей, математическая и прикладная статистика: учебник / Ю. С. Харин, Н. М. Зуев, Е. Е. Жук. – Мн. : БГУ, 2011.
6. Миленький, А. В. Классификация сигналов в условиях неопределенности / А. В. Миленький. – М. : Сов. радио, 1975.
7. Дуда, Р. Распознавание образов и анализ сцен / Р. Дуда, П. Харт. М. : Мир, 1976.
8. Харин, Ю. С. Математические и компьютерные основы статистического моделирования и анализа – Мн. : БГУ, 2008.
9. Харин, Ю. С. Компьютерный учебник по статистике данных / Ю. С. Харин, М. С. Абрамович, В. И. Малюгин. – Мн. : БГУ, 2001.
10. Харин, Ю. С. Практикум на ЭВМ по математической статистике / Ю. С. Харин, М. Д. Степанова. – Мн. : Университетское, 1987.

Перечень дополнительной литературы

1. Патрик, Э. Основы теории распознавания образов / Э. Патрик. – М. : Мир, 1980.
2. Прикладная статистика. Основы моделирования и первичная обработка данных / С. А. Айвазян [и др.]. – М. : Финансы и статистика, 1983.
3. Прикладная статистика. Исследование зависимостей / С. А. Айвазян [и др.]. – М. : Финансы и статистика, 1985.
4. Харин, Ю. С. Робастность в статистическом распознавании образов / Ю. С. Харин. – Мн. : Университетское, 1992.
5. Мандель, И. Д. Кластерный анализ / И. Д. Мандель. – М. : Финансы и статистика, 1988.
6. Жук, Е. Е. Устойчивость в кластер-анализе многомерных наблюдений / Е. Е. Жук, Ю. С. Харин. – Мн. : БГУ, 1998.
7. Вапник, В. Н. Теория распознавания образов / В. Н. Вапник, А. Я. Червоненкис. – М. : Наука, 1974.
8. Боровиков, В. П. STATISTICA. Статистический анализ и обработка данных в среде Windows / В. П. Боровиков. – М. : Филинь, 1997.

9. Ивченко, Г. И. Введение в математическую статистику / Г. И. Ивченко, Ю. И. Медведев. – М. : URSS, 2009.

Примерный перечень тем для коллоквиумов

1. Параметрический дискриминантный анализ
2. Формирование информативных признаков методом главных компонент.

Методические рекомендации по организации самостоятельной работы обучающихся

Для организации самостоятельной работы студентов по учебной дисциплине следует использовать современные информационные технологии: разместить в сетевом доступе комплекс учебных и учебно-методических материалов (учебно-программные материалы, ссылки на учебные издания для теоретического изучения дисциплины, методические указания к лабораторным занятиям, материалы текущего контроля и текущей аттестации, позволяющие определить соответствие учебной деятельности обучающихся требованиям образовательных стандартов высшего образования и учебно-программной документации, в т.ч. вопросы для подготовки к зачету, задания, тесты, вопросы для самоконтроля, тематика рефератов и др., список рекомендуемой литературы, информационных ресурсов и др.). Эффективность самостоятельной работы студентов проверяется в ходе текущего и итогового контроля знаний. Для общей оценки качества усвоения студентами учебного материала рекомендуется использование рейтинговой системы.

Перечень рекомендуемых средств диагностики

Для текущего контроля качества усвоения знаний студентами используется следующий диагностический инструментарий:

1. Устная форма: устные опросы; защиты отчетов по домашним заданиям, при выполнении студентами лабораторных работ; проведение коллоквиума; защита подготовленного студентом реферата (рефераты используются для обобщения и систематизации учебного материала; в процессе подготовки реферата студент мобилизует и актуализирует имеющиеся умения, приобретает самостоятельно новые знания, необходимые для раскрытия темы, сопоставляя разные позиции и точки зрения).
2. Письменная форма: письменные контрольные работы по отдельным темам учебной дисциплины.

Методика формирования итоговой оценки

Формой текущей аттестации по учебной дисциплине «Распознавание образов» учебным планом предусмотрены зачет и экзамен.

При оценивании реферата внимание обращается на:

- содержание, корректность и последовательность изложения – 35%;
- релевантность и полноту раскрытие темы – 20 %;
- самостоятельность суждений – 35%;
- оформление – 10%.

Рекомендуется использовать рейтинговую оценку знаний студента, дающую возможность проследить и оценить динамику процесса достижения целей обучения. Рейтинговая оценка предусматривает использование весовых коэффициентов для текущего контроля знаний и текущей аттестации студентов по дисциплине. Примерные весовые коэффициенты, определяющие вклад текущего контроля знаний в рейтинговую оценку:

- подготовка реферата – 15 %;
- работа на лабораторных занятиях – 35 %;
- контрольные работы – 25 %;
- коллоквиум – 25 %.

Итоговая оценка формируется на основе:

- 1) Правил проведения аттестации студентов (Постановление Министерства образования Республики Беларусь № 53 от 29 мая 2012г.);
- 2) Положение о рейтинговой системе оценки знаний по дисциплине в БГУ (Приказ ректора БГУ от 18.08.2015 № 382-ОД);
- 3) Критериев оценки знаний студентов (письмо Министерства образования от 22.12.2003).

ПРОТОКОЛ СОГЛАСОВАНИЯ УЧЕБНОЙ ПРОГРАММЫ УВО

Название учебной дисциплины, с которой требуется согласование	Название кафедры	Предложения об изменениях в содержании учебной программы учреждения высшего образования по учебной дисциплине	Решение, принятое кафедрой, разработавшей учебную программу (с указанием даты и номера протокола)
Имитационное и статистическое моделирование	Математического моделирования и анализа данных	нет	Оставить содержание учебной дисциплины без изменения, протокол № 13 от 29 марта 2018.
Модели и методы теории прогнозирования	Математического моделирования и анализа данных	нет	Оставить содержание учебной дисциплины без изменения, протокол № 13 от 29 марта 2018.

ДОПОЛНЕНИЯ И ИЗМЕНЕНИЯ К УЧЕБНОЙ ПРОГРАММЕ ПО ИЗУЧАЕМОЙ УЧЕБНОЙ ДИСЦИПЛИНЕ

на ____ / ____ учебный год

№ п/п	Дополнения и изменения	Основание

Учебная программа пересмотрена и одобрена на заседании кафедры
_____ (протокол № ____ от _____ 20__ г.)

Заведующий кафедрой

УТВЕРЖДАЮ

Декан факультета
