

Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Новосибирский государственный технический университет»

Кафедра теоретической и прикладной информатики

“УТВЕРЖДАЮ”
ДЕКАН ФПМИ
д.т.н. Тимофеев В. С.
“ ____ ” _____ г.

РАБОЧАЯ ПРОГРАММА УЧЕБНОЙ ДИСЦИПЛИНЫ
Основы теории машинного обучения

Образовательная программа: 02.04.03 Математическое обеспечение и администрирование информационных систем, магистерская программа: Математическое и программное обеспечение информационных технологий

Курс: 1, семестр : 2

Факультет прикладной математики и информатики,

		Семестр
№	Вид деятельности	2
1	Всего зачетных единиц (кредитов)	3
2	Всего часов	108
3	Всего занятий в контактной форме, час.	71
4	Лекции, час.	18
5	Практические занятия, час.	0
6	Лабораторные занятия, час.	36
7	из них в активной и интерактивной форме, час.	30
8	Аттестация, час.	2
9	Консультации, час.	15
10	Самостоятельная работа, час.	37
11	Виды самостоятельной работы (курсовой проект, курсовая работа, РГЗ, подготовка к контрольной работе)	
12	Вид аттестации	ДЗ

Рабочая программа составлена на основании федерального государственного образовательного стандарта высшего образования по направлению (специальности): 02.04.03 Математическое обеспечение и администрирование информационных систем

ФГОС введен в действие приказом №1416 от 30.10.2014 г. , дата утверждения: 26.11.2014 г.

Место дисциплины в структуре учебного плана: Блок 1, вариативная

Рабочая программа разработана на основе компетентностной модели выпускника по направлению (специальности): 02.04.03 Математическое обеспечение и администрирование информационных систем

Рабочая программа обсуждена на заседании кафедры

ТПИ, протокол заседания кафедры №4 от 20.06.2017

Утверждена на совете факультета прикладной математики и информатики, протокол № 6 от 21.06.2017

Программу разработал:

профессор, д.т.н. Попов А. А.

Заведующий кафедрой:

доцент, д.т.н. Чубич В. М.

Ответственный за образовательную программу:

заведующий кафедрой Чубич В. М.

1. Внешние требования

Таблица 1.1

Компетенция ФГОС: ОПК.10 владение навыками использования основных моделей информационных технологий и способов их применения для решения задач в предметных областях; в части следующих результатов обучения:	
у5. Уметь применять современные методы построения зависимостей в условиях структурной и параметрической неопределенности по неоднородным данным, решать задачи классификации с использованием метрических алгоритмов, метода опорных векторов, деревьев решений, алгоритмов усиления, решать задачи поиска ассоциаций и кластеризации	
Компетенция ФГОС: ОПК.4 владение теоретическими основами информатики как науки; знание проблем современной информатики, ее категории и связи с другими научными дисциплинами, понимание основных этапов и тенденции развития программирования, математического обеспечения и информационных технологий; в части следующих результатов обучения:	
з4. Знать основные задачи машинного обучения по прецедентам (построение зависимостей, классификация, кластеризация, поиск ассоциаций, визуализация многомерных данных) и методы их решения.	
у4. Уметь использовать базовые знания линейной алгебры, вычислительной математики, теории вероятности и математической статистики, статистических методов анализа данных, методов оптимизации при разработке вычислительных алгоритмов машинного обучения	
Компетенция ФГОС: ОПК.9 владение навыками разработки моделирующих алгоритмов и реализации их на базе языков и пакетов прикладных программ моделирования; в части следующих результатов обучения:	
у3. Владеть основами программирования в одном из современных математических пакетов статистической направленности	
у4. Уметь реализовывать вычислительные алгоритмы построения зависимостей по экспериментальным данным, решения задач классификации, кластеризации	
Компетенция ФГОС: ПК.1 владение навыками применения математических основ информатики при разработке и исследовании нового программного обеспечения; в части следующих результатов обучения:	
з1. Уметь реализовывать различные алгоритмы, используемые при решении различных задач по распознаванию образов	

2. Требования НГТУ к результатам освоения дисциплины

Таблица 2.1

Результаты изучения дисциплины по уровням освоения (иметь представление, знать, уметь, владеть)	Формы организации занятий
<i>Основы теории машинного обучения</i>	
ОПК.4.з4 Знать основные задачи машинного обучения по прецедентам (построение зависимостей, классификация, кластеризация, поиск ассоциаций, визуализация многомерных данных) и методы их решения.	
1. Знать основные задачи машинного обучения по прецедентам (построение зависимостей, классификация, кластеризация, поиск ассоциаций) и методы их решения.	Лекции; Самостоятельная работа
ОПК.4.у4 Уметь использовать базовые знания линейной алгебры, вычислительной математики, теории вероятности и математической статистики, статистических методов анализа данных, методов оптимизации при разработке вычислительных алгоритмов машинного обучения	
2. Уметь использовать базовые знания линейной алгебры, вычислительной математики, теории вероятности и математической статистики, статистических методов анализа данных, методов оптимизации при разработке вычислительных алгоритмов машинного обучения	Лекции; Самостоятельная работа
ОПК.9.у3 Владеть основами программирования в одном из современных математических пакетов статистической направленности	

3. Владеть основами программирования в одном из современных математических пакетов статистической направленности	Лекции; Лабораторные работы; Самостоятельная работа
ОПК.9.y4 Уметь реализовывать вычислительные алгоритмы построения зависимостей по экспериментальным данным, решения задач классификации, кластеризации	
4. Уметь реализовывать вычислительные алгоритмы построения зависимостей по экспериментальным данным, решения задач классификации	Лекции; Лабораторные работы
ОПК.10.y5 Уметь применять современные методы построения зависимостей в условиях структурной и параметрической неопределенности по неоднородным данным, решать задачи классификации с использованием метрических алгоритмов, метода опорных векторов, деревьев решений, алгоритмов усиления, решать задачи поиска ассоциаций и кластеризации	
5. Уметь применять современные методы построения зависимостей в условиях структурной и параметрической неопределенности по неоднородным данным, решать задачи классификации с использованием метода опорных векторов, решать задачи поиска ассоциаций	Лекции; Лабораторные работы
ПК.1.31 Уметь реализовывать различные алгоритмы, используемые при решении различных задач по распознаванию образов	
6. Уметь реализовывать различные алгоритмы, используемые при решении различных задач по распознаванию образов	Лекции; Лабораторные работы

3. Содержание и структура учебной дисциплины

Таблица 3.1

Темы лекций	Активные формы, час.	Часы	Ссылки на результаты обучения	Учебная деятельность
Семестр: 2				
Дидактическая единица: Дисперсионный анализ данных.				
1. Линейные модели с качественными факторами. Основы теории обобщенных обратных матриц. Решение нормальных уравнений с использованием обобщенного обращения. Параметрические функции, допускающие оценку (ФДО). Определение и оценивание ФДО. Теорема Гаусса-Маркова. Проверка гипотез. Редуцирование моделей	0	4	1, 2, 3, 4, 5	Учебная деятельность состоит в изучении теоретического материала, подготовки и выполнения лабораторных работ, их защита
Дидактическая единица: Построение робастных регрессионных моделей				
2. Понятие устойчивого оценивания параметров линейных регрессионных моделей. Виды функций потерь. Итерационный метод наименьших квадратов.	0	2	1, 2, 3, 4, 5	Учебная деятельность состоит в изучении теоретического материала, подготовки и выполнения лабораторных работ, их защита
Дидактическая единица: Построение зависимостей в рамках концепции нечетких систем				
3. Система нечетких правил Такаги-Сугено. Оценивание параметров. Выбор моделей оптимальной сложности. Методы нечеткой кластеризации для построения зависимостей. Построение робастных решений.	0	4	1, 2, 3, 4, 5	Учебная деятельность состоит в изучении теоретического материала, подготовки и выполнения лабораторных работ, их защита
Дидактическая единица: Построение регрессионных зависимостей по методу опорных векторов				

4. Функция потерь Вапника. Квадратичная функция потерь. Прямая и двойственная задача квадратичного программирования. Построение робастных решений. Построение разреженных решений	0	4	1, 2, 3, 4, 5	Учебная деятельность состоит в изучении теоретического материала, подготовки и выполнения лабораторных работ, их защита
Дидактическая единица: Методы опорных векторов для задач классификации				
5. Оптимальная разделяющая гиперплоскость. Понятие зазора между классами (margin). Случаи линейной разделимости и отсутствия линейной разделимости. Связь с минимизацией регуляризованного эмпирического риска. Кусочно-линейная функция потерь. Задача квадратичного программирования и двойственная задача. Понятие опорных векторов. Рекомендации по выбору константы C. Функция ядра (kernel functions), спрямляющее пространство, теорема Мерсера. Способы конструктивного построения ядер. Примеры ядер. Метод релевантных векторов RVM	0	2	1, 2, 3, 4, 5, 6	Учебная деятельность состоит в изучении теоретического материала, подготовки и выполнения лабораторных работ, их защита
Дидактическая единица: Поиск ассоциативных правил				
6. Понятие ассоциативного правила и его связь с понятием логической закономерности. Примеры прикладных задач: анализ рыночных корзин, выделение терминов и тематики текстов. Алгоритм APriori. Два этапа: поиск частых наборов и рекурсивное порождение ассоциативных правил. Недостатки и пути усовершенствования алгоритма APriori.	0	2	1, 3, 4, 5	Учебная деятельность состоит в изучении теоретического материала, подготовки и выполнения лабораторных работ, их защита

Таблица 3.2

Темы лабораторных работ	Активные формы, час.	Часы	Ссылки на результаты обучения	Учебная деятельность
Семестр: 2				
Дидактическая единица: Дисперсионный анализ данных.				

1. Построение и анализ линейных моделей с номинальными переменными	5	6	3, 4	Приобретает практические навыки по решению задачи построения и анализа линейных моделей с номинальными переменными с использованием редукции исходной модели
Дидактическая единица: Построение робастных регрессионных моделей				
2. Построение робастных регрессионных моделей.	5	6	3, 4	Приобретает практические навыки по решению задачи построения робастных регрессионных моделей
Дидактическая единица: Построение зависимостей в рамках концепции нечетких систем				
3. Построение зависимостей в рамках концепции нечетких систем	5	6	4	Решает задачу построения регрессионных моделей в рамках концепции нечетких систем.
Дидактическая единица: Построение регрессионных зависимостей по методу опорных векторов				
4. Построение регрессионных зависимостей по методу опорных векторов	5	6	4	Учебная деятельность состоит в изучении теоретического материала, подготовки и выполнения лабораторных работ, их защита
Дидактическая единица: Методы опорных векторов для задач классификации				
5. Решение задачи классификации по методу опорных векторов	5	6	4, 5, 6	Приобретает практические навыки по решению задачи классификации с использованием метода опорных векторов
Дидактическая единица: Поиск ассоциативных правил				
6. Поиск ассоциативных правил в наборах данных	5	6	3, 4	Приобретает практические навыки по решению задачи поиска ассоциативных правил в наборах данных

4. Самостоятельная работа обучающегося

№	Виды самостоятельной работы	Ссылки на результаты обучения	Часы на выполнение	Часы на консультации
Семестр: 2				
1	Подготовка к занятиям	2, 3	27	12
Подготовка и изучения материалов по теме лабораторных работ: Попов А. А. Основы теории машинного обучения: методические указания к лабораторным работам для студентов ФПМИ, обучающихся по направлению 02.04.03 Математическое обеспечение и администрирование информационных систем [Электронный ресурс] : учебно-методическое пособие / А. А. Попов ; Новосиб. гос. техн. ун-т. - Новосибирск, [2017]. - Режим доступа: http://elibrary.nstu.ru/source?bib_id=vtls000234571 . - Загл. с экрана.				
2	Подготовка к аттестации	1, 2	10	3
Подготовка к аттестации на основе повторения материала по дисциплине: Организация самостоятельной работы студентов Новосибирского государственного технического университета : методическое руководство / Новосиб. гос. техн. ун-т ; [сост.: Ю. В. Никитин, Т. Ю. Сурнина]. - Новосибирск, 2016. - 19, [1] с. : табл.. - Режим доступа: http://elibrary.nstu.ru/source?bib_id=vtls000234042				

5. Технология обучения

Для организации и контроля самостоятельной работы обучающихся, а также проведения консультаций применяются информационно-коммуникационные технологии (табл. 5.1).

Таблица 5.1

Деятельность	Информационно-коммуникационные технологии
Информирование	e-mail; Личный типовой сайт
Консультирование	e-mail; Личный типовой сайт
Контроль	e-mail
Размещение учебных материалов	Личный типовой сайт

Таблица 5.2

Активные и интерактивные формы проведения занятий

№	Наименование активных форм
1	Проблемный метод
Краткое описание применения: Обсуждение постановки задачи, вопросов реализации алгоритмов ее решения, полученных результатов	

6. Правила аттестации обучающихся по учебной дисциплине

Для аттестации обучающихся по дисциплине используется балльно-рейтинговая система (БРС), позволяющая выставять оценки по традиционной шкале и 15-уровневой ECTS. Краткая информация о БРС приведена в табл. 6.1.

Таблица 6.1

Оцениваемые виды деятельности обучающихся	Максимальный балл
Семестр: 2	
<i>Лабораторная №1:</i>	12
Контролирующие материалы - список вопросов	
<i>Лабораторная №2:</i>	12
Контролирующие материалы - список вопросов	
<i>Лабораторная №3:</i>	14
Контролирующие материалы - список вопросов	
<i>Лабораторная №4:</i>	14
Контролирующие материалы - список вопросов	
<i>Лабораторная №5:</i>	14
Контролирующие материалы - список вопросов	
<i>Лабораторная №6:</i>	14
Контролирующие материалы - список вопросов	
<i>Зачет:</i>	20
Контролирующие материалы - список вопросов	

В таблице 6.2 представлено соответствие форм контроля заявляемым требованиям к результатам освоения дисциплины.

Таблица 6.2

Коды компетенций ФГОС	Результаты обучения	Формы контроля	
			Зачет
ОПК.10	у5. Уметь применять современные методы построения зависимостей в условиях структурной и параметрической неопределенности по неоднородным данным, решать задачи классификации с использованием метрических алгоритмов, метода опорных векторов, деревьев решений, алгоритмов усиления, решать задачи поиска ассоциаций и кластеризации		+
ОПК.4	з4. Знать основные задачи машинного обучения по прецедентам (построение зависимостей, классификация, кластеризация, поиск ассоциаций, визуализация многомерных данных) и методы их решения.		+
	у4. Уметь использовать базовые знания линейной алгебры, вычислительной математики, теории вероятности и математической статистики, статистических методов анализа данных, методов оптимизации при разработке вычислительных алгоритмов машинного обучения		+
ОПК.9	у3. Владеть основами программирования в одном из современных математических пакетов статистической направленности		+
	у4. Уметь реализовывать вычислительные алгоритмы построения зависимостей по экспериментальным данным, решения задач классификации, кластеризации		+
ПК.1	з1. Уметь реализовывать различные алгоритмы, используемые при решении различных задач по распознаванию образов		+

Фонд оценочных средств по дисциплине представлен в приложении № 1 к рабочей программе.

7. Литература

Основная литература

1. Попов А. А. Оптимальное планирование эксперимента в задачах структурной и параметрической идентификации моделей многофакторных систем : [монография] / А. А. Попов. - Новосибирск, 2013. - 295 с. : ил., табл.. - Парал. тит. л. англ.
2. Friedman J. The Elements of Statistical Learning [electronic resource] : : Data Mining, Inference, and Prediction // by Jerome Friedman, Robert Tibshirani, Trevor Hastie. - New York, NY :, 2009. : v.: digital // Springer eBooks. - Режим доступа: <http://dx.doi.org/10.1007/b94608>
3. Vapnik V. Estimation of Dependences Based on Empirical Data [electronic resource] : : Empirical Inference Science Afterword of 2006 // by Vladimir Vapnik. - New York, NY :, 2006. : v.: digital // Springer e-books. - Режим доступа: <http://dx.doi.org/10.1007/0-387-34239-7>
4. Clarke B. Principles and Theory for Data Mining and Machine Learning [electronic resource] // by Bertrand Clarke, Ernest Fokoue, Hao Helen Zhang. - New York, NY :, 2009. : v.: digital // Springer eBooks. - Режим доступа: <http://dx.doi.org/10.1007/978-0-387-98135-2>
5. Bull L. Learning Classifier Systems in Data Mining [electronic resource] // edited by Larry Bull, Ester Bernado-Mansilla, John Holmes. - Berlin, Heidelberg :, 2008. : v.: digital // Springer eBooks. - Режим доступа: <http://dx.doi.org/10.1007/978-3-540-78979-6>

6. Гулятьева Т. А. Методы статистического обучения в задачах регрессии и классификации : [монография] / Т. А. Гулятьева, А. А. Попов, А. С. Саутин. - Новосибирск, 2016. - 321, [1] с. : ил., табл.. - Режим доступа: http://elibrary.nstu.ru/source?bib_id=vtls000227465. - Парал. тит. л. и огл. англ..

Дополнительная литература

1. Прикладная статистика. Классификация и снижение размерности : справочное издание / С. А. Айвазян, В. М. Бухштабер, И. С. Енюков, Л. Д. Мешалкин ; под ред. С. А. Айвазяна. - М., 1989. - 606, [1] с. : ил.. - Перед вып. дан. 1-й авт.: Сергей Арутюнович Айвазян.
2. Загоруйко Н. Г. Прикладные методы анализа данных и знаний / Н. Г. Загоруйко. - Новосибирск, 1999. - 269 с. : ил.. - Библиогр.: с. 247-260. - Предм. указ.: с. 261-263.
3. Попов А. А. Конструирование линейных регрессионных моделей с разнотипными переменными : Учеб. пособие для IV курса ФПМИ дневной формы обучения / А. А. Попов; Новосиб. гос. техн. ун-т. - Новосибирск, 1999. - 55 с. : ил.
4. Чубукова И. А. Data Mining : учебное пособие / И. А. Чубукова. - М., 2006. - 382 с. : ил.
5. Шлезингер М. И. Десять лекций по статистическому и структурному распознаванию / М. Шлезингер, В. Главач. - Киев, 2004. - 535 с. : ил.
6. Desu M. M. Nonparametric Statistical Methods for Complete and Censored Data / M. M. Desu and D. Raghavarao. - Boca Raton, 2004. - XIV, 367 p. : ill.. - Пер. загл.: Непараметрические методы статистики для полных и цензурированных данных.
7. Малов С. В. Регрессионный анализ. Теоретические основы и практические рекомендации / С. В. Малов ; С.-Петербург. гос. ун-т. - Санкт-Петербург, 2013. - 275 с.
8. Воскобойников Ю. Е. Регрессионный анализ данных в пакете Mathcad : учебное пособие / Ю. Е. Воскобойников. - Санкт-Петербург [и др.], 2011. - 223, [1] с. : ил., табл. + 1 CD.

Интернет-ресурсы

1. ЭБС НГТУ : <http://elibrary.nstu.ru/>
2. ЭБС «Издательство Лань» : <https://e.lanbook.com/>
3. ЭБС IPRbooks : <http://www.iprbookshop.ru/>
4. ЭБС "Znaniy.com" : <http://znaniy.com/>
5. :

8. Методическое и программное обеспечение

8.1 Методическое обеспечение

1. Попов А. А. Основы теории машинного обучения: методические указания к лабораторным работам для студентов ФПМИ, обучающихся по направлению 02.04.03 Математическое обеспечение и администрирование информационных систем [Электронный ресурс] : учебно-методическое пособие / А. А. Попов ; Новосиб. гос. техн. ун-т. - Новосибирск, [2017]. - Режим доступа: http://elibrary.nstu.ru/source?bib_id=vtls000234571. - Загл. с экрана.
2. Организация самостоятельной работы студентов Новосибирского государственного технического университета : методическое руководство / Новосиб. гос. техн. ун-т ; [сост.: Ю. В. Никитин, Т. Ю. Сурнина]. - Новосибирск, 2016. - 19, [1] с. : табл.. - Режим доступа: http://elibrary.nstu.ru/source?bib_id=vtls000234042

8.2 Специализированное программное обеспечение

1 Python

9. Материально-техническое обеспечение

Презентационное оборудование

№	Наименование	Назначение
1	Презентационное оборудование (мультимедиа-проектор, экран, компьютер для управления)	Для проведения лекций

Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Новосибирский государственный технический университет»

Кафедра теоретической и прикладной информатики

“УТВЕРЖДАЮ”
ДЕКАН ФПИ
д.т.н., доцент В.С. Тимофеев
“ ” _____ Г.

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

УЧЕБНОЙ ДИСЦИПЛИНЫ

Основы теории машинного обучения

Образовательная программа: 02.04.03 Математическое обеспечение и администрирование информационных систем, магистерская программа: Математическое и программное обеспечение информационных технологий

1. Обобщенная структура фонда оценочных средств учебной дисциплины

Обобщенная структура фонда оценочных средств по дисциплине Основы теории машинного обучения приведена в Таблице.

Таблица

Формируемые компетенции	Показатели сформированности компетенций (знания, умения, навыки)	Темы	Этапы оценки компетенций	
			Мероприятия текущего контроля (курсовой проект, РГЗ(Р) и др.)	Промежуточная аттестация (экзамен, зачет)
ОПК.10 владение навыками использования основных моделей информационных технологий и способов их применения для решения задач в предметных областях	у5. Уметь применять современные методы построения зависимостей в условиях структурной и параметрической неопределенности по неоднородным данным, решать задачи классификации с использованием метрических алгоритмов, метода опорных векторов, деревьев решений, алгоритмов усиления, решать задачи поиска ассоциаций и кластеризации	Решение задачи классификации по методу опорных векторов		Зачет, вопросы 1-38
ОПК.4 владение теоретическими основами информатики как науки; знание проблем современной информатики, ее категории и связи с другими научными дисциплинами, понимание основных этапов и тенденции развития программирования, математического обеспечения и информационных технологий	з4. Знать основные задачи машинного обучения по прецедентам (построение зависимостей, классификация, кластеризация, поиск ассоциаций, визуализация многомерных данных) и методы их решения.	Линейные модели с качественными факторами. Основы теории обобщенных обратных матриц. Решение нормальных уравнений с использованием обобщенного обращения. Параметрические функции, допускающие оценку (ФДО). Определение и оценивание ФДО. Теорема Гаусса-Маркова. Проверка гипотез. Редуцирование моделей Оптимальная разделяющая гиперплоскость. Понятие зазора между классами (margin). Случаи линейной разделимости и отсутствия линейной разделимости. Связь с минимизацией регуляризованного эмпирического риска. Кусочно-линейная функция потерь. Задача квадратичного программирования и двойственная задача. Понятие опорных векторов. Рекомендации по выбору константы C. Функция ядра (kernel functions),		Зачет, вопросы 1-38

		спрямляющее пространство, теорема Мерсера. Способы конструктивного построения ядер. Примеры ядер. Метод релевантных векторов RVM Понятие ассоциативного правила и его связь с понятием логической закономерности. Примеры прикладных задач: анализ рыночных корзин, выделение терминов и тематики текстов. Алгоритм APriori. Два этапа: поиск частых наборов и рекурсивное порождение ассоциативных правил. Недостатки и пути усовершенствования алгоритма APriori. Понятие устойчивого оценивания параметров линейных регрессионных моделей. Виды функций потерь. Итерационный метод наименьших квадратов. Система нечетких правил Такаги-Сугено. Оценивание параметров. Выбор моделей оптимальной сложности. Методы нечеткой кластеризации для построения зависимостей. Построение робастных решений. Функция потерь Вапника. Квадратичная функция потерь. Прямая и двойственная задача квадратичного программирования. Построение робастных решений. Построение разреженных решений		
ОПК.4	у4. Уметь использовать базовые знания линейной алгебры, вычислительной математики, теории вероятности и математической статистики, статистических методов анализа данных, методов оптимизации при разработке вычислительных алгоритмов машинного обучения	Линейные модели с качественными факторами. Основы теории обобщенных обратных матриц. Решение нормальных уравнений с использованием обобщенного обращения. Параметрические функции, допускающие оценку (ФДО). Определение и оценивание ФДО. Теорема Гаусса-Маркова. Проверка гипотез. Редуцирование моделей Оптимальная разделяющая гиперплоскость. Понятие зазора между классами (margin). Случаи линейной разделимости и отсутствия линейной разделимости. Связь с минимизацией регуляризованного эмпирического риска. Кусочно-линейная функция потерь. Задача квадратичного программирования и двойственная задача. Понятие опорных векторов.		Зачет, вопросы 1-38

		Рекомендации по выбору константы C. Функция ядра (kernel functions), спрямляющее пространство, теорема Мерсера. Способы конструктивного построения ядер. Примеры ядер. Метод релевантных векторов RVM Понятие устойчивого оценивания параметров линейных регрессионных моделей. Виды функций потерь. Итерационный метод наименьших квадратов. Система нечетких правил Такаги-Сугено. Оценивание параметров. Выбор моделей оптимальной сложности. Методы нечеткой кластеризации для построения зависимостей. Построение робастных решений. Функция потерь Вапника. Квадратичная функция потерь. Прямая и двойственная задача квадратичного программирования. Построение робастных решений. Построение разреженных решений		
ОПК.9 владение навыками разработки моделирующих алгоритмов и реализации их на базе языков и пакетов прикладных программ моделирования	у3. Владеть основами программирования в одном из современных математических пакетов статистической направленности	Построение и анализ линейных моделей с номинальными переменными		Зачет, вопросы 1-38
ОПК.9	у4. Уметь реализовывать вычислительные алгоритмы построения зависимостей по экспериментальным данным, решения задач классификации, кластеризации	Поиск ассоциативных правил в наборах данных Построение зависимостей в рамках концепции нечетких систем Построение и анализ линейных моделей с номинальными переменными Построение робастных регрессионных моделей.		Зачет, вопросы 1-38
ПК.1/НИ владение навыками применения математических основ информатики при разработке и исследовании нового программного обеспечения	з1. Уметь реализовывать различные алгоритмы, используемые при решении различных задач по распознаванию образов	Оптимальная разделяющая гиперплоскость. Понятие зазора между классами (margin). Случаи линейной разделимости и отсутствия линейной разделимости. Связь с минимизацией регуляризованного эмпирического риска. Кусочно-линейная функция потерь. Задача квадратичного программирования и двойственная задача. Понятие опорных векторов.		Зачет, вопросы 1-38

		Рекомендации по выбору константы C . Функция ядра (kernel functions), спрямляющее пространство, теорема Мерсера. Способы конструктивного построения ядер. Примеры ядер. Метод релевантных векторов RVM. Решение задачи классификации по методу опорных векторов		
--	--	---	--	--

2. Методика оценки этапов формирования компетенций в рамках дисциплины.

Промежуточная аттестация по дисциплине проводится в 2 семестре - в форме дифференцированного зачета, который направлен на оценку сформированности компетенций ОПК.10, ОПК.4, ОПК.9, ПК.1/НИ.

Зачет проводится в письменной форме, по билетам. Время подготовки на зачете составляет 1,5 часа.

На основании приведенных далее критериев можно сделать общий вывод о сформированности компетенций ОПК.10, ОПК.4, ОПК.9, ПК.1/НИ за которые отвечает дисциплина, на разных уровнях.

Общая характеристика уровней освоения компетенций.

Ниже порогового. Уровень выполнения работ не отвечает большинству основных требований, теоретическое содержание курса освоено частично, пробелы могут носить существенный характер, необходимые практические навыки работы с освоенным материалом сформированы не достаточно, большинство предусмотренных программой обучения учебных заданий не выполнены или выполнены с существенными ошибками.

Пороговый. Уровень выполнения работ отвечает большинству основных требований, теоретическое содержание курса освоено частично, но пробелы не носят существенного характера, необходимые практические навыки работы с освоенным материалом в основном сформированы, большинство предусмотренных программой обучения учебных заданий выполнено, некоторые виды заданий выполнены с ошибками.

Базовый. Уровень выполнения работ отвечает всем основным требованиям, теоретическое содержание курса освоено полностью, без пробелов, некоторые практические навыки работы с освоенным материалом сформированы недостаточно, все предусмотренные программой обучения учебные задания выполнены, качество выполнения ни одного из них не оценено минимальным числом баллов, некоторые из выполненных заданий, возможно, содержат ошибки.

Продвинутый. Уровень выполнения работ отвечает всем требованиям, теоретическое содержание курса освоено полностью, без пробелов, необходимые практические навыки работы с освоенным материалом сформированы, все предусмотренные программой обучения учебные задания выполнены, качество их выполнения оценено числом баллов, близким к максимальному.

Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Новосибирский государственный технический университет»

Кафедра теоретической и прикладной информатики

Паспорт зачета

По дисциплине Основы теории машинного обучения, семестр 2

1. Методика оценки

Зачет проводится в письменной форме, по билетам. В ходе зачета преподаватель вправе задавать студенту дополнительные вопросы из общего перечня (п. 4).

Форма билета для зачета

Дисциплина Основы теории машинного обучения
БИЛЕТ № 1

№	Вопрос	Баллов
1	Дисперсионный анализ. Классификация по одному признаку	4
2	Робастная регрессия. М-оценки	4
3	Основные понятия нечетких множеств. Функции принадлежности, их виды	4
4	Построение зависимостей по методу SVM	4
5	Поиск ассоциаций	4
	Итого	20

Составитель д.т.н., проф. Попов А.А.

_____ (подпись)

«___» _____ 20__ г.

Заведующий кафедрой д.т.н. В.М. Чубич

_____ (подпись)

«___» _____ 20__ г.

2. Критерии оценки

- Ответ считается **неудовлетворительным**, если оценка составляет от 0 до 6 баллов, Студент при ответе на вопросы не дает определений основных понятий;
- Ответ засчитывается на **пороговом** уровне, если оценка составляет от 7 до 10 баллов. Студент при ответе на вопросы дает определение основных понятий, знает базовые алгоритмы, при решении;
- Ответ засчитывается на **базовом** уровне, если оценка составляет от 11 до 15 баллов. Студент при ответе на вопросы формулирует основные понятия, методы, алгоритмы, проводит анализ причин, условий;
- Ответ засчитывается на **продвинутом** уровне, если оценка составляет от 16 до 20 баллов. Студент владеет математическим аппаратом, методами прикладной статистики, базовыми алгоритмами машинного обучения.

Кроме того, сформированность компетенций проверяется при проведении мероприятий текущего контроля, указанных в таблице раздела 1.

Общие правила выставления оценки по дисциплине определяются балльно-рейтинговой системой, приведенной в рабочей программе дисциплины.

3. Шкала оценки

Итоговая аттестация по курсу проводится в виде зачета. Для получения допуска к зачету студент должен, как правило, выполнить и защитить все лабораторные работы. Основным критерий допуска к экзамену – число набранных баллов в течение семестра, которое должно составить не менее 43 баллов.

Максимальное количество набираемых баллов в период итоговой аттестации равно 20 баллам. Если студент по результатам итоговой аттестации набирает **менее 7 баллов**, то ему выставляется оценка "неудовлетворительно" уровня FX вне зависимости от числа набранных баллов в семестре с возможностью пересдачи. При успешной пересдаче ему выставляется оценка "удовлетворительно" уровня E. Общее количество набранных баллов по дисциплине определяется простым суммированием набранных баллов по лабораторным работам и на зачете.

Перевод набранного количества баллов в 15 бальную систему и 4-х бальную систему осуществляется по следующей схеме:

ECTS	A+	A	A-	B+	B	B-	C+	C	C-	D+	D	D-	E	FX	F
Баллы	96-100	93-95	90-92	87-89	83-86	80-82	76-79	73-75	70-72	66-69	63-65	60-62	50-59	25-49	0-24
	отлично			хорошо			удовлетворительно						неудовлетв.		

4. Вопросы к зачету по дисциплине «Основы теории машинного обучения»

Полный перечень вопросов
по дисциплине "Основы теории машинного обучения"

1. Постановка задач обучения по прецедентам. Объекты и признаки. Типы шкал: бинарные, номинальные, порядковые, количественные.
2. Типы задач: классификация, регрессия, прогнозирование, кластеризация. Примеры прикладных задач.
3. Основы теории обобщенных обратных матриц. Типы обобщенных обратных матриц и их свойства.
4. Свойства псевдообратных матриц.
5. Вычисление обобщенных обратных матриц: алгоритм, основанный на обычном обращении матриц, алгоритм, основанный на использовании элементарных операторов.
6. Вычисление псевдообратных матриц: алгоритм Гревилля.
7. Решение СЛАУ с использованием обобщенного обращения. Параметрические функции, допускающие оценку (ФДО). Определение и оценивание ФДО.
8. Дисперсия оценок ФДО, свойства остаточной суммы квадратов.
9. Распределения основных статистик. Проверка линейных гипотез.
10. Базис ФДО для линейной модели ДА.
11. Базис ФДО для модели ДА с взаимодействиями уровней факторов.
12. Базис ФДО для линейной модели с порядковыми переменными.
13. Оценивание параметров и ФДО в редуцированной модели.
14. МНК и устойчивость
15. Робастные оценки, М-оценки
16. Итеративный МНК
17. Метод псевдонаблюдений
18. Виды функций потерь, требования к ним
19. Построение зависимостей по методу LS SVM
20. Задача квадратичного программирования и двойственная задача. Понятие опорных векторов.
21. Способы конструктивного построения ядер. Примеры ядер.
22. Функция потерь Вапника.
23. Формулировка задачи оптимизации в терминах двойственных переменных
24. Классификация по методу LS SVM
25. Настройка параметров алгоритма LS SVM.
26. Оптимальная разделяющая гиперплоскость.
27. Понятие зазора между классами (margin).
28. Случаи линейной разделимости и отсутствия линейной разделимости.
29. Лингвистические переменные
30. Функции принадлежности

31. Фаззификация и дефаззификация
32. Модели Такаги-Сугено
33. Локальное и глобальное оценивание параметров TS моделей
34. Критерии выбора модели оптимальной сложности
35. Алгоритмы нечеткой кластеризации в задаче построения TS моделей.
36. Понятие ассоциативного правила и его связь с понятием логической закономерности.
37. Понятия поддержки, достоверности;
38. Алгоритм APriori и его разновидности.