

Министерство науки и высшего образования Российской Федерации  
Федеральное государственное автономное образовательное учреждение  
высшего образования «Новосибирский национальный исследовательский  
государственный университет» (Новосибирский государственный университет, НГУ)

**Факультет информационных технологий**

Согласовано

Декан ФИТ НГУ

  
М.М. Лаврентьев

«28» сентября 2022 г.

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ**

**МЕТОДЫ ОБЪЯСНИМОГО ИСКУССТВЕННОГО ИНТЕЛЛЕКТА**

Научная специальность: 2.3.5 Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей

Направленность (профиль): Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей

Форма обучения: очная

Разработчик:

доцент кафедры общей информатики ФИТ,  
кандидат физико-математических наук



Г.Э. Яхьяева

Заведующий кафедрой общей информатики ФИТ,  
доктор физико-математических наук



Д.Е. Пальчунов

Руководитель программы:

заведующий кафедрой систем информатики ФИТ,  
доктор физико-математических наук



М.М. Лаврентьев

## **Аннотация к рабочей программе дисциплины «Методы объяснимого искусственного интеллекта»**

Дисциплина «Методы объяснимого искусственного интеллекта» реализуется в рамках программы аспирантуры по научной специальности 2.3.5 «Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей» и направленности (профилю): Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей по очной форме обучения на русском языке.

Дисциплина «Методы объяснимого искусственного интеллекта» входит в блок элективных дисциплин, реализуемых в рамках программы аспирантуры и направлена на формирование знаний и умений, связанных с разработками в области объяснимого (доверенного) искусственного интеллекта.

**Целью курса** является изучение методов и алгоритмов объяснимого искусственного интеллекта («белого ящика»), а также методов интерпретации работы алгоритмов «черного ящика». В рамках данного курса аспиранты будут знакомиться с научными публикациями (не более пятилетней давности) из ведущих журналов по данному направлению. На занятиях эти публикации будут реферироваться и анализироваться. В результате успешного освоения курса аспиранты изучат новейшие результаты в области объяснимого (доверенного) искусственного интеллекта, а также приобретут навык работы с научной литературой. Дисциплина рассчитана на два семестра.

**Перечень основных разделов дисциплины:** технологии инженерии знаний, методы представления знаний, онтологическое моделирование, анализ формальных понятий, основные понятия теории нечетких множеств, нечеткие модели, методы нечеткого моделирования, экспертные системы.

**Общий объем дисциплины** – 4 зачетных единицы (144 часа). Из которых лекции составляют 64 часа, самостоятельная работа 74 часа, консультации 2 часа.

**Правила аттестации по дисциплине.** Промежуточная аттестация по дисциплине «Методы объяснимого искусственного интеллекта» проводится в виде сдачи дифференцированного зачета. Дифференцированный зачет проводится в устной форме, по билетам. Билет выбирается обучающимся случайным образом. При подготовке ответа на вопросы билета не разрешается использование каких-либо источников информации.

## 1. Результаты освоения дисциплины:

- Знать формальные методы представления знаний, методы автоматизации дедуктивных, индуктивных рассуждений и рассуждений по аналогии, основы онтологического моделирования,
- Уметь формулировать профессиональные задачи в терминах инженерии знаний, делать сравнительный анализ и обосновать выбор модели и средства представления знаний, строить модель заданной предметной области с использованием изученных средств представления знаний.

## 2. Трудоемкость дисциплины по видам учебной деятельности

Трудоемкость дисциплины – 4 з.е. (144 ч)

Форма промежуточной аттестации: диф.зачет.

| №  | Вид деятельности   | Количество часов |           |
|----|--|------------------|-----------|
|    |  | 3 семестр        | 4 семестр |
| 1  | Лекции, час.   | 32               | 32        |
| 2  | Практические занятия, час.   |                  |           |
| 3  | Лабораторные занятия, час  |                  |           |
| 4  | Занятий в контактной форме без учета промежуточной аттестации, час, из них | 34               | 32        |
| 5  | в электронной форме, час.  |                  |           |
| 6  | аудиторных занятий, час.   | 32               | 32        |
| 7  | из них в активной и интерактивной форме, час.                              |                  |           |
| 8  | консультаций, час.   | 2                |           |
| 9  | Самостоятельная работа, час.   | 36               | 38        |
| 10 | Всего, ч   | 72               | 72        |

## 3. Содержание дисциплины

Лекции 3 семестр, 32 часа

| Наименование тем и их содержание   | Объем час |
|--|-----------|
| 1. <b>Технологии инженерии знаний:</b> История развития систем искусственного интеллекта. Методы извлечения знаний. Этапы построения базы знаний.  | 8         |
| 2. <b>Методы представления знаний:</b> процедурные представления, логические представления, семантические сети, фреймы, системы продукций. Интегрированные методы представления знаний. Языки представления знаний. Базы знаний. | 8         |
| 3. <b>Онтологическое моделирование:</b> Понятие, определение, онтология. Онтологии предметных областей. Разработка и применение онтологий. Логика описаний. Семантический Веб.   | 8         |
| 4. <b>Анализ формальных понятий:</b> Формальные контексты и решетки формальных понятий. Методы построения решеток формальных понятий. Свойства решеток формальных понятий. Импликативные зависимости и ассоциативные правила.    | 8         |

Лекции 4 семестр, 32 часа

| Содержание практического занятия   | Объем час |
|--|-----------|
| <b>1. Основные понятия теории нечетких множеств.</b> Историческая справка. Методология гранулированного подхода. Методология лингвистических рассуждений.  | <b>6</b>  |
| <b>2. Нечеткие модели, методы нечеткого моделирования.</b> Нечеткие модели. Свойства правил, баз правил и нечетких моделей. Сокращение баз правил. Масштабирование входов и выходов нечеткой модели. Экстраполяция в нечетких моделях. Типы нечетких моделей.  | <b>6</b>  |
| <b>3. Экспертные системы.</b> Статические и динамические экспертные системы. Составные части экспертной системы. Участники процесса проектирования экспертной системы. Этапы проектирования экспертной системы: идентификация, концептуализация, формализация, реализация, тестирование, опытная эксплуатация. | <b>8</b>  |
| <b>4. Вопросно-ответные системы.</b> История области применения ВОС. ВОС с открытым и закрытым доменом. Архитектура ВОС. Виды вопросов. Обработка вопросов. Косвенные вопросы. Проблемы формулировки ответов.  | <b>6</b>  |
| <b>5. Рекомендательные системы.</b> Коллаборативная фильтрация. Контентная фильтрация. Проблема узких рекомендаций. Проблема холодного старта. Разреженность данных.   | <b>6</b>  |

Самостоятельная работа аспиранта 3 семестр, 36 часов

| Перечень занятий на СР аспиранта  | Объем час |
|---|-----------|
| 1. Формальные языки и способы их описания. Классификация формальных грамматик. Их использование в лексическом и синтаксическом анализе. | 8         |
| 2. Информационно-поисковые системы. Классификация. Методы реализации и ускорения поиска.  | 10        |
| 3. Алгоритмы классификации и кластеризации.   | 10        |
| 4. Алгоритмы обработки временных рядов и поиска аномалий  | 8         |

Самостоятельная работа аспиранта 4 семестр, 38 часов

| Перечень занятий на СР аспиранта  | Объем час |
|---|-----------|
| 1. Алгоритмы обработки текста. Представления текста для последующей обработки (bag of words, tf-idf, word embeddings, one-hot encoding) | 8         |
| 2. Генетические алгоритмы. Роевые алгоритмы.  | 10        |
| 3. Нечеткое моделирование, вероятностное моделирование  | 10        |
| 4. Алгоритмы интерпретации нейронных сетей  | 10        |

#### **4. Перечень учебно-методических материалов, необходимых для изучения дисциплины (модуля)**

1. Павлов, С.И. Системы искусственного интеллекта : учебное пособие / С.И. Павлов. - Томск : Томский государственный университет систем управления и радиоэлектроники, 2011. - Ч. 1. - 175 с. - ISBN 978-5-4332-0013-5 ; [Электронный ресурс]. - URL:

<http://biblioclub.ru/index.php?page=book&id=208933>

#### **5. Перечень учебных изданий, необходимых для освоения дисциплины**

2. Павлов, С.И. Системы искусственного интеллекта : учебное пособие / С.И. Павлов. - Томск : Томский государственный университет систем управления и радиоэлектроники, 2011. - Ч. 2. - 194 с. - ISBN 978-5-4332-0014-2 ; [Электронный ресурс]. - URL:

<http://biblioclub.ru/index.php?page=book&id=208939>

#### **6. Материально-техническая база, необходимая для осуществления образовательного процесса по дисциплине**

Для реализации дисциплины «Методы объяснимого искусственного интеллекта» используются специальные помещения:

1. Учебные аудитории для проведения занятий лекционного типа, занятий семинарского типа, курсового проектирования (выполнения курсовых работ), групповых и индивидуальных консультаций, текущего контроля, промежуточной и итоговой аттестации;
2. Помещения для самостоятельной работы обучающихся;
3. Лаборатории;
4. Помещения для хранения и профилактического обслуживания учебного оборудования.

Учебные аудитории укомплектованы специализированной мебелью и техническими средствами обучения, служащими для представления учебной информации большой аудитории.

Помещения для самостоятельной работы обучающихся оснащены компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду НГУ

Для проведения занятий лекционного типа предлагаются следующие наборы демонстрационного оборудования и учебно-наглядных пособий:

- комплект лекций-презентаций по темам дисциплины;

Материально-техническое обеспечение образовательного процесса по дисциплине для аспирантов из числа лиц с ограниченными возможностями здоровья осуществляется согласно «Порядку организации и осуществления образовательной деятельности по образовательным программам для инвалидов и лиц с ограниченными возможностями здоровья в Новосибирском государственном университете»

## **7. Оценочные средства для проведения промежуточной аттестации обучающихся по дисциплине (модулю)**

### **7.1 Порядок проведения текущего контроля и промежуточной аттестации по дисциплине**

#### ***Промежуточная аттестация:***

Промежуточная аттестация по дисциплине «Методы объяснимого искусственного интеллекта» проводится в виде сдачи дифференцированного зачета.

Дифференцированный зачет проводится в устной форме, по билетам. Билет выбирается обучающимся случайным образом. При подготовке ответа на вопросы билета не разрешается использование каких-либо источников информации. В процессе ответа обучающегося на вопросы билета преподаватель может задавать дополнительные вопросы по темам дисциплины. Результаты промежуточной аттестации определяются оценками «отлично», «хорошо», «удовлетворительно», «неудовлетворительно». Оценки «отлично», «хорошо», «удовлетворительно» означают успешное прохождение промежуточной аттестации.

#### ***Описание критериев и шкал оценивания результатов освоения дисциплины:***

| <b>Результат освоения дисциплины</b>   | <b>Критерии оценивания результатов освоения дисциплины</b>  | <b>Шкала оценивания</b> |
|--|---|-------------------------|
| <ul style="list-style-type: none"> <li>- Знать формальные методы представления знаний, методы автоматизации дедуктивных, индуктивных рассуждений и рассуждений по аналогии, основы онтологического моделирования,</li> <li>- Уметь формулировать профессиональные задачи в терминах инженерии знаний, делать сравнительный анализ и обосновать выбор модели и средства представления знаний, строить модель заданной предметной</li> </ul> | <p>Демонстрирует четкое представление о формальных методах представления знаний. Способен в полной мере формализовывать знания; строить модель знаний. Демонстрирует полное владение основными принципами инженерии знаний.</p> | Отлично                 |
| <ul style="list-style-type: none"> <li>- Знать формальные методы представления знаний, методы автоматизации дедуктивных, индуктивных рассуждений и рассуждений по аналогии, основы онтологического моделирования,</li> <li>- Уметь формулировать профессиональные задачи в терминах инженерии знаний, делать сравнительный анализ и обосновать выбор модели и средства представления знаний, строить модель заданной предметной</li> </ul> | <p>Демонстрирует знание формальных методов представления знаний. Способен в достаточной мере формализовывать знания; строить модель знаний. Демонстрирует</p>   | Хорошо                  |

|  |   |                     |
|--|---|---------------------|
| области с использованием изученных средств представления знаний. | достаточное владение основными принципами инженерии знаний.   |                     |
|  | Имеет неполное представление о формальных методах представления знаний. Демонстрирует слабое умение формализовывать знания; строить модель знаний. Демонстрирует слабое владение основными принципами инженерии знаний. | удовлетворительно   |
|  | Не имеет представление о формальных методах представления знаний. Не способен формализовывать знания, строить модель знаний. Демонстрирует непонимание основных принципов инженерии знаний.                             | неудовлетворительно |

Результаты промежуточной аттестации, проводимой в форме дифференцированного зачета, определяются оценками «отлично», «хорошо», «удовлетворительно», «неудовлетворительно». Оценки «отлично», «хорошо», «удовлетворительно» означают успешное прохождение промежуточной аттестации.

***Типовые контрольные задания и иные материалы, необходимые для оценки результатов освоения дисциплины (оценочные материалы)***

Перечень вопросов к дифференцируемому зачету:

1. Область искусственного интеллекта (ИИ).
2. Основные понятия и определения ИИ.
3. Краткий исторический обзор развития работ в области ИИ.
4. Функциональная структура систем искусственного интеллекта (СИИ).
5. Определение процесса решения проблемной задачи.
6. Модели задач, их классификация.
7. Человеко-машинные системы решения сложных задач.
8. Интеллектуальный интерфейс и его структура в современных ЭВМ.
9. Роль знаний в процессе решения задач.
10. Формальное представление задачи.
11. Анализ условий задачи для выбора методов решения.
12. Решение задач методом поиска в пространстве состояний.
13. Решение задач методом редукции.
14. Метод ключевых состояний и ключевых операторов.
15. Метод анализа средств и целей.

16. Решение задач методом дедуктивного вывода.
17. Логические системы.
18. Метод резолюции и его применение для решения задач.
19. Модели мира и их роль в решении задач.
20. Формальные модели представления знаний.
21. Данные и знания.
22. Переход от Базы Данных к Базе Знаний.
23. Продукционные системы.
24. Представление простых фактов в логических системах.
25. Семантические сети, фреймы, сценарии.
26. Представления знаний для структурированных объектов, инженерия знаний.
27. Базы знаний.
28. Понятия обучение и самообучения.
29. Классификация методов обучения.
30. Принципы индуктивного обучения для детерминированных и нечетких систем.
31. Обучение с использованием генетических алгоритмов.
32. Нейронные сети и используемые для них методы обучения.