**Санкт-Петербургский государственный университет**

**Р А Б О Ч А Я П Р О Г Р А М М А**

**УЧЕБНОЙ ДИСЦИПЛИНЫ**

Интеллектуальные системы

Intelligent Systems

**Язык(и) обучения**

русский

Трудоемкость в зачетных единицах: 3

Регистрационный номер рабочей программы: 003691

**Раздел 1. Характеристики учебных занятий**

**1.1. Цели и задачи учебных занятий**

Цель изучения дисциплины: ознакомление обучающихся с историей и современными подходами к интеллектуальному анализу данных; изучение основ машинного обучения и различных специализированных пакетов; общение обучающихся с экспертами, применяющими техники машинного обучения и анализа данных в промышленных масштабах; решения обучающимися задач находящихся на границе возможностей современной науки и выходящие за пределы; выступление обучающихся с докладами перед коллегами и приглашенными экспертами.

**1.2. Требования подготовленности обучающегося к освоению содержания учебных занятий (пререквизиты)**

Программа курса предназначена для обучающихся 4 курса и рассчитана на слушателей, изучавших математический и дискретный анализ, алгебру и теорию вероятности в программе обучения первых 6-ти семестров.

Максимальная эффективность программы будет обеспечена при условии, что обучающийся:

• знаком с теорией вероятности и математической статистикой;  
• владеет навыками поиска научной информации и инструментов в сети Интернет;  
• способен к написанию программ на языках Java и Python;  
• не потерял желание учится новому и делать то, что до него не делал никто.

**1.3. Перечень результатов обучения (learning outcomes)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| № | Наименование категории (группы) компетенций | Код и наименование компетенции | Планируемые результаты обучения, обеспечивающие формирование компетенции | Код индикатора и индикатор достижения универсальной компетенции |
|  | 1 | 2 | 3 | 4 |
| 1 | Общепрофессиональные компетенции | ОПК-1 – способен применять естественнонаучные и общеинженерные знания, методы математического анализа и моделирования, теоретического и экспериментального исследования в профессиональной деятельности; | Знание ключевых фактов и общее представление об истории и современных методов машинного обучения и анализа данных | ОПК-1.1 Уметь идентифицировать возможные проблемы и пути их решения |
| 2 | Общепрофессиональные компетенции | ОПК-2 Способен понимать принципы работы современных информационных технологий и программных средств, в том числе отечественного производства, и использовать их при решении задач профессиональной деятельности | Закрепление навыков использования математического анализа и теории вероятности при решении актуальных практических задач | ОПК-2.2 Уметь проводить формализацию и алгоритмизацию поставленных задач |
| 3 | Профессиональные компетенции | ПКП-1 – способен проектировать программные системы; | Умение всесторонне анализировать поставленные задачи, включая поиск подходов, методов и инструментов для решения близких задач | ПКП-1.1 Уметь разрабатывать технические спецификации на программные компоненты и их взаимодействие |
| 4 | Профессиональные компетенции | ПКП-2 – способен использовать основные модели информационных технологий и способы их применения для решения задач в предметных областях; | Умение всесторонне анализировать поставленные задачи, включая поиск подходов, методов и инструментов для решения близких задач | ПКП-2.1 Уметь описывать алгоритмы компонентов, включая методы и схемы |
| 5 | Профессиональные компетенции | ПКП-3 – способен разрабатывать моделирующие алгоритмы и реализации их на базе языков и пакетов прикладных программ моделирования; | Умение всесторонне анализировать поставленные задачи, включая поиск подходов, методов и инструментов для решения близких задач | ПКП-3.1 Разработка модели бизнес-процессов заказчика |
| 6 | Профессиональные компетенции | ПКП-7 – способен систематизировать и применять знания о содержании основных этапов и тенденций развития программирования, математического обеспечения и информационных технологий; | Умение всесторонне анализировать поставленные задачи, включая поиск подходов, методов и инструментов для решения близких задач | ПКП-7.1 Умеет проводить технические и управленческие ревизии создаваемого программного средства |
| 7 | Универсальные компетенции | УК-1 – способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач; | Навыки командной научно-исследовательской работы над решением актуальных задач машинного обучения и анализа данных | УК 1.2. Определяет информацию, необходимую для решения поставленной задачи; |
| 8 | Универсальные компетенции | УК-4 – способен осуществлять деловую коммуникацию в устной и письменной формах на государственном языке Российской Федерации и иностранном(ых) языке(ах); | Опыт презентации своих результатов и подготовки обзорных докладов | УК-4.3. Ведет деловую переписку на русском языке с учетом особенностей стилистики официальных и неофициальных писем; |
| 9 | Универсальные компетенции | УКБ-3 – способен понимать сущность и значение информации в развитии общества, использовать основные методы получения и работы с информацией с учетом современных технологий цифровой экономики и информационной безопасности; | Опыт презентации своих результатов и подготовки обзорных докладов | УКБ-3.2. Точно определяет тип и форму необходимой информации. |

**1.4. Перечень и объём активных и интерактивных форм учебных занятий**

Практические занятия – 10 ак. ч.

**Раздел 2. Организация, структура и содержание учебных занятий**

**2.1. Организация учебных занятий**

**2.1.1 Основной курс**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Трудоёмкость, объёмы учебной работы и наполняемость групп обучающихся | | | | | | | | | | | | | | | | | | |
| Код модуля в составе дисциплины,  практики и т.п. | Контактная работа обучающихся с преподавателем | | | | | | | | | | | | Самостоятельная работа | | | | Объём активных и интерактивных  форм учебных занятий | Трудоёмкость |
| лекции | семинары | консультации | практические  занятия | лабораторные работы | контрольные работы | коллоквиумы | текущий контроль | промежуточная  аттестация | итоговая аттестация | под руководством преподавателя | в присутствии  преподавателя | сам. раб. с использованием  методических материалов | текущий контроль (сам.раб.) | промежуточная аттестация (сам.раб.) | итоговая аттестация  (сам.раб.) |
| ОСНОВНАЯ ТРАЕКТОРИЯ | | | | | | | | | | | | | | | | | | |
| Форма обучения: очная | | | | | | | | | | | | | | | | | | |
| Семестр 7 | 30 |  | 2 | 30 |  |  |  |  | 2 |  |  |  | 23 |  | 21 |  | 10 | 3 |
|  | 2-42 |  | 2-42 | 2-25 |  |  |  |  | 2-25 |  |  |  | 1-1 |  | 1-1 |  |  |  |
| ИТОГО | 30 |  | 2 | 30 |  |  |  |  | 2 |  |  |  | 23 |  | 21 |  |  | 3 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Виды, формы и сроки текущего контроля успеваемости и промежуточной аттестации | | | | | | |
| Код модуля в составе дисциплины, практики и т.п. | Формы текущего контроля успеваемости | | Виды промежуточной аттестации | | Виды итоговой аттестации  (только для программ итоговой аттестации и дополнительных образовательных программ) | |
| Формы | Сроки | Виды | Сроки | Виды | Сроки |
| ОСНОВНАЯ ТРАЕКТОРИЯ | | | | | | |
| Форма обучения: очная | | | | | | |
| Семестр 7 |  |  | экзамен, устно, традиционная форма | по графику промежуточной аттестации |  |  |

**2.2. Структура и содержание учебных занятий**

1. Введение в курс и исторический экскурс:
   1. понятия интеллекта и критерии интеллектуальности, некоторые базовые понятия системного анализа;
   2. обзор задач, решаемых с помощью ЭВМ, и обладающих свойствами интеллектуальности;
   3. основные методы используемые при решении таких задач;
   4. важные точки роста и нерешенные задачи.
2. Основы машинного обучения и анализа данных:
   1. основные задачи машинного обучения и краткая история развития;
   2. обучение с учителем: задачи аппроксимации, прогнозирования и классификации;
   3. обучение без учителя: задачи кластеризации и поиска характеристических признаков;
   4. анализ данных: поиск корреляций и закономерностей;
   5. визуализация исходных данных и результатов работы алгоритмов;
   6. проверка качества работы алгоритмов машинного обучения и анализа данных;
   7. особенности разработки в проектах, связанных с машинным обучением;
   8. обзор популярных инструментов;
3. Рекомендательные системы:
   1. область применения рекомендательных систем, основные задачи и вехи развития;
   2. коллаборативные рекомендательные системы: ближайшие соседи, факторизация матриц, случайные блуждания;
   3. социальные рекомендательные системы: применимость и особенности использования;
   4. контентные рекомендательные системы: решение задач холодного старта;
   5. комплексные методы: ансамбли рекомендательных систем, обучение в реальном времени;
   6. методы оценки качества работы рекомендательных систем;
   7. существующие инструменты разработки рекомендательных систем.
4. Анализ естественных языков:
   1. история развития подходов к анализу текстов на естественных языках: от грамматик Хомского к машинному обучению и корпоральной лингвистике;
   2. текущие актуальные задачи анализа и синтеза текстов;
   3. базовые техники разбора текста: определение языка, токенизация, выделение предложений, определение частей речи, поиск именованных сущностей;
   4. унификация термов: стемминг, лематизация, исправление опечаток;
   5. анализ семантики текста: латентно семантический анализ и вероятностные генеративные модели;
   6. анализ эмоциональной окраски текстов: словарные методы и машинное обучение;
   7. методы оценки качества на разных этапах анализа;
   8. существующие инструменты для анализа текстов на естественных языках;
5. Анализ мультимедийного контента:
   1. задачи, связанны с анализом мультимедийного контента (ауди, фото и видео материалы);
   2. краткий обзор используемого математического аппарата;
   3. аудио контент: подходы к идентификации, дедупликации и извлечению характеристических признаков;
   4. фото контент: подходы к идентификации, дедупликации и извлечению характеристических признаков;
   5. видео контент: подходы к идентификации, дедупликации и извлечению характеристических признаков;
   6. особенности построения систем на основе локально-чувствительных хешей;
   7. методы оценки качества в задачах анализа мультимедийного контента;
   8. существующие инструменты для анализа мультимедийного контента.
6. Анализ больших данных:
   1. история появления термина «большие данные» и варианты текущего его толкования;
   2. особенности связанные с применением алгоритмов машинного обучения к большим данным;
   3. виды анализа данных: операционный, интерактивный, потоковый и пакетный;
   4. Apache Hadoop: история развития, текущий статус, возможности и область применения;
   5. некоторые продукты Hadoop-экосистемы: Spark, Pig, Kafka, Samza, Accumulo;
   6. некоторые инструменты машинного обучения, масштабируемые на большие данные: MLLib, GraphX, Mahaout, Mr.LDA.

**Раздел 3. Обеспечение учебных занятий**

**3.1. Методическое обеспечение**

**3.1.1 Методические указания по освоению дисциплины**

Методические материалы включают в себя следующие типы материалов — учебники, учебные пособия, методические указания для обучающихся, Интернет-ресурсы, электронные учебные пособия, с опорой на которые проводится аудиторная работа.

**3.1.2 Методическое обеспечение самостоятельной работы**

Самостоятельная работа обучающегося, как вид деятельности, стимулирующий активность, самостоятельность, познавательный интерес с целью поиска необходимой информации, приобретения знаний, использования этих знаний для решения учебных, научных и профессиональных задач, представляет собой важную составляющую учебного процесса, которой отводится не менее половины учебного времени при очной форме обучения. Время, отводимое на самостоятельную работу, должно использоваться обучающимися для наиболее полного освоения учебной дисциплины. Следовательно, организация эффективной внеаудиторной самостоятельной работы в процессе обучения требует, с одной стороны, создание условий, призванных обеспечить рациональное и планомерное управление учебной деятельностью, протекающей в отсутствие преподавателя, и тщательной подготовки целого ряда учебных пособий, снабженных методическими указаниями, с другой стороны.

К числу методических пособий относятся общие методические рекомендации и указания по самостоятельной работе.

Роль преподавателя в организации самостоятельной работы состоит в координации действий обучающихся в освоении дисциплины, в методическом и организационном обеспечении учебного процесса. Взаимодействие между преподавателем и обучающимся осуществляется в форме консультаций. Преподаватели также оказывают помощь обучающимся по планированию и организации самостоятельной работы.

**3.1.3 Методика проведения текущего контроля успеваемости и промежуточной аттестации и критерии оценивания**

***3.1.3.1. Методика проведения текущего контроля успеваемости и промежуточной аттестации***

Текущая успеваемость контролируется в рамках решения практических задач по блокам курса (4 задачи на семестр). Обучающиеся, успешно сдавшие решения всех 4-х задач, допускаются к экзамену, который проводится в формате устного собеседования. На собеседовании необходимо продемонстрировать свои знания как минимум в 3 различных модулях.

**3.1.3.2. Критерии оценивания итогового процента освоения дисциплины**

Практические задачи оцениваются по шкале от 0 (решение отсутствует или обладает существенными недостатками) до 1 (решение существенными недостатками не обладает). Результирующая оценка находится в диапазоне от 0 до 100 и вычисляется по формуле MAX(0, (n/N – 0.6)) \* 2.5 \* 100, где n – суммарный балл, набранный обучающимся на момент аттестации, N – максимально возможный суммарный балл за курс. Например, обучающийся, успешно сдавший 60% заданий, получает 0 баллов, 80% – 50, 100% – 100 баллов.

На теоретическом экзамене ответ на каждый вопрос про конкретный модуль и на дополнительные вопросы оценивается по шкале от 0 (нет ответа) до 10 (очень хороший ответ), далее оценка усредняется. Результат переводится в диапазон от 0 до 100.

Итоговый процент выполнения целей изучения дисциплины вычисляется как минимум из оценки за практические задачи и оценки за теоретический экзамен. Далее применяется следующее правило выставления оценки:

|  |  |  |
| --- | --- | --- |
| Итоговый процент  выполнения, % | Оценка СПбГУ при  проведении экзамена | Оценка  ECTS |
| 90-100 | отлично | A |
| 80-89 | хорошо | B |
| 70-79 | хорошо | C |
| 61-69 | удовлетворительно | D |
| 50-60 | удовлетворительно | E |
| менее 50 | неудовлетворительно | F |

**3.1.4 Методические материалы для проведения текущего контроля успеваемости и промежуточной аттестации (контрольно-измерительные материалы, оценочные средства)**

Аппарат контроля за усвоением материала включает в себя экзамен по итогам курса.

Пример списка вопросов, выносимых на экзамен:

* 1. Обучение с учителем: задачи аппроксимации,
  2. Обучение с учителем: задачи прогнозирования;
  3. Обучение с учителем: задачи классификации;
  4. Обучение без учителя: задачи кластеризации;
  5. Обучение без учителя: задачи поиска характеристических признаков;
  6. Анализ данных: поиск корреляций и закономерностей;
  7. Визуализация данных;
  8. Метрики оценки качества работы алгоритмов машинного обучения и анализа данных;
  9. Область применения рекомендательных систем, основные задачи;
  10. Коллаборативные рекомендательные системы: ближайшие соседи, факторизация матриц, случайные блуждания;
  11. Социальные рекомендательные системы: применимость и особенности использования;
  12. Контентные рекомендательные системы: решение задач холодного старта;
  13. Комплексные методы: ансамбли рекомендательных систем, обучение в реальном времени;
  14. Методы оценки качества работы рекомендательных систем;
  15. Существующие инструменты разработки рекомендательных систем.
  16. Базовые техники разбора текста: определение языка, токенизация;
  17. Базовые техники разбора текста: выделение предложений, определение частей речи;
  18. Базовые техники разбора текста: поиск именованных сущностей;
  19. Унификация термов: стемминг, лематизация, исправление опечаток;
  20. Анализ семантики текста: латентно семантический анализ и вероятностные генеративные модели;
  21. Анализ эмоциональной окраски текстов: словарные методы и машинное обучение;
  22. Методы оценки качества на разных этапах анализа текста;
  23. Аудио контент: подходы к идентификации, дедупликации и извлечению характеристических признаков;
  24. Фото контент: подходы к идентификации, дедупликации и извлечению характеристических признаков;
  25. Видео контент: подходы к идентификации, дедупликации и извлечению характеристических признаков;
  26. Особенности построения систем на основе локально-чувствительных хешей;
  27. Методы оценки качества в задачах анализа мультимедийного контента;
  28. Существующие инструменты для анализа мультимедийного контента.
  29. Виды анализа данных: операционный, интерактивный, потоковый и пакетный;
  30. Apache hadoop: возможности и область применения;

|  |  |  |
| --- | --- | --- |
| № | Код индикатора и индикатор достижения универсальной компетенции | Контрольно-измерительные материалы (КИМ) (тестовые вопросы, контрольные задания, кейсы и пр.) |
|  | 1 | 2 |
| 1 | ОПК-1.1 Уметь идентифицировать возможные проблемы и пути их решения | в качестве уровня сформированности каждой компетенции используется итоговые баллы за задачи в диапазоне от 0 до 100 |
| 2 | ОПК-2.2 Уметь проводить формализацию и алгоритмизацию поставленных задач | в качестве уровня сформированности каждой компетенции используется итоговые баллы за задачи в диапазоне от 0 до 100 |
| 3 | ПКП-1.1 Уметь разрабатывать технические спецификации на программные компоненты и их взаимодействие | в качестве уровня сформированности каждой компетенции используется итоговые баллы за задачи в диапазоне от 0 до 100 |
| 4 | ПКП-2.1 Уметь описывать алгоритмы компонентов, включая методы и схемы | в качестве уровня сформированности каждой компетенции используется итоговые баллы за задачи в диапазоне от 0 до 100 |
| 5 | ПКП-3.1 Разработка модели бизнес-процессов заказчика | в качестве уровня сформированности каждой компетенции используется итоговые баллы за задачи в диапазоне от 0 до 100 |
| 6 | ПКП-7.1 Умеет проводить технические и управленческие ревизии создаваемого программного средства | ответы вопросы экзамена и на дополнительные вопросы оцениваются по шкале от 0 (нет ответа) до 10 (очень хорошо), далее оценка усредняется и переводится в диапазон от 0 до 100. Если за задачи набрано 100 баллов, то компетенции сформированы на 100 баллов |
| 7 | УК 1.2. Определяет информацию, необходимую для решения поставленной задачи; | ответы вопросы экзамена и на дополнительные вопросы оцениваются по шкале от 0 (нет ответа) до 10 (очень хорошо), далее оценка усредняется и переводится в диапазон от 0 до 100. Если за задачи набрано 100 баллов, то компетенции сформированы на 100 баллов |
| 8 | УК-4.3. Ведет деловую переписку на русском языке с учетом особенностей стилистики официальных и неофициальных писем; | ответы вопросы экзамена и на дополнительные вопросы оцениваются по шкале от 0 (нет ответа) до 10 (очень хорошо), далее оценка усредняется и переводится в диапазон от 0 до 100. Если за задачи набрано 100 баллов, то компетенции сформированы на 100 баллов |
| 9 | УКБ-3.2. Точно определяет тип и форму необходимой информации. | ответы вопросы экзамена и на дополнительные вопросы оцениваются по шкале от 0 (нет ответа) до 10 (очень хорошо), далее оценка усредняется и переводится в диапазон от 0 до 100. Если за задачи набрано 100 баллов, то компетенции сформированы на 100 баллов |

**3.1.5 Методические материалы для оценки обучающимися содержания и качества учебного процесса**

Для оценки обучающимися содержания и качества учебного процесса применяется анкетирование в соответствии с методикой и графиком, утвержденными в установленном порядке.

**3.2. Кадровое обеспечение**

**3.2.1 Образование и (или) квалификация штатных преподавателей и иных лиц, допущенных к проведению учебных занятий**

К преподаванию дисциплины могут быть допущены преподаватели, имеющие диплом о высшем образовании по соответствующему направлению.

**3.2.2 Обеспечение учебно-вспомогательным и (или) иным персоналом**

Специальных требований нет.

**3.3. Материально-техническое обеспечение**

**3.3.1 Характеристики аудиторий (помещений, мест) для проведения занятий**

Учебные аудитории для проведения учебных занятий, оснащенные стандартным оборудованием, используемым для обучения в СПбГУ в соответствии с требованиями материально-технического обеспечения.

**3.3.2 Характеристики аудиторного оборудования, в том числе неспециализированного компьютерного оборудования и программного обеспечения общего пользования**

Стандартное оборудование, используемое для обучения в СПбГУ. MS Windows, MS Office, Mozilla FireFox, Google Chrome, Acrobat Reader DC, WinZip, Антивирус Касперского

**3.3.3 Характеристики специализированного оборудования**

Специальных требований нет.

**3.3.4 Характеристики специализированного программного обеспечения**

На каждом компьютере для обучающихся должны быть установлены следующие программные системы:

Интернет браузер (например: Firefox, Safari, MS Internet Explorer, Google Chrome)

и возможность выхода в Internet.

**3.3.5 Перечень и объёмы требуемых расходных материалов**

Специальных требований нет.

**3.4. Информационное обеспечение**

**3.4.1 Список литературы**

1.Смолин, Денис Валерьевич . Введение в искусственный интеллект : конспект лекций / Д. В. Смолин. - 2-е изд., перераб. - М. : Физматлит, 2007. - 259 с.   
<https://proxy.library.spbu.ru:2385/reader/book/2325/#1>

2.Тулупьев, Александр Львович . Байесовские сети доверия : логико-вероятностный вывод в ациклических направленных графах / А. Л. Тулупьев, А. В. Сироткин, С. И. Николенко ; Санкт-Петербургский государственный университет, Российская академия наук. Санкт-Петербургский институт информатики и автоматизации. - СПб. : Издательство Санкт-Петербургского университета, 2009. - 399 с.

3. Пегат, Анджей. Нечеткое моделирование и управление: Пер. с англ. / А. Пегат; пер. А. Г. Подвесовский, пер., ред. пер. Ю. В. Тюменцев. - М.: БИНОМ. Лаборатория знаний, 2009. - 798 с.

**3.4.2 Перечень иных информационных источников**

• Сайт Научной библиотеки им. М. Горького СПбГУ: <http://www.library.spbu.ru/>

• Электронный каталог Научной библиотеки им. М. Горького СПбГУ: <http://www.library.spbu.ru/cgi-bin/irbis64r/cgiirbis_64.exe?C21COM=F&I21DBN=IBIS&P21DBN=IBIS>

• Перечень электронных ресурсов, находящихся в доступе СПбГУ: <http://cufts.library.spbu.ru/CRDB/SPBGU/>

• Перечень ЭБС, на платформах которых представлены российские учебники, находящиеся в доступе СПбГУ: <http://cufts.library.spbu.ru/CRDB/SPBGU/browse?name=rures&resource%20type=8>

**Раздел 4. Разработчики программы**

Азимов Рустам Шухратуллович, магистр, ассистент кафедры информационно-аналитических систем, [st013567@student.spbu.ru](mailto:st013567@student.spbu.ru)