

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Сибирский государственный индустриальный университет»
Кафедра прикладных информационных технологий и программирования

УТВЕРЖДАЮ
Проректор по учебной и
воспитательной работе
_____ М.В. Темлянцев
подпись
« ____ » _____ 20__ г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Машинное обучение для задач управления

09.04.03 «Прикладная информатика»
(направленность (профиль): «Системы корпоративного управления»)

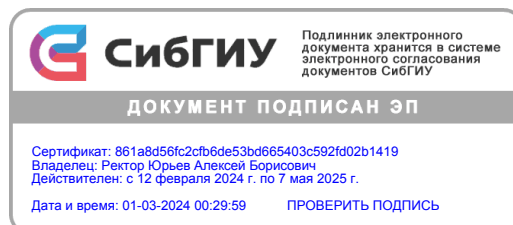
Квалификация выпускника
Магистр

Форма обучения
Очная форма

Срок обучения: 2 года

Год начала подготовки 2022

Новокузнецк
2022



1 Цели и задачи освоения учебной дисциплины

Целями учебной дисциплины являются:

- формирование у обучающихся знаний и навыков применения методов машинного обучения для задач управления; моделирования и прогнозирования ML-методами, а также оценки построенной модели.

Задачами учебной дисциплины являются:

- изучение методов сбора и подготовки реальных данных для машинного обучения;
- изучение этапов проектирования признаков для построения моделей машинного обучения;
- изучение процесса построения моделей классификации и регрессии методами ML;
- изучение подходов к оценке построенных моделей.

2 Место учебной дисциплины в структуре ООП по направлению подготовки (специальности)

Учебная дисциплина относится к учебным дисциплинам части, формируемой участниками образовательных отношений **Блока 1 «Дисциплины (модули)»** ООП по направлению подготовки (специальности) 09.04.03 «Прикладная информатика».

Учебная дисциплина базируется на предварительном усвоении обучающимися учебных дисциплин:

- Python для анализа данных;
- Методология научного познания.

Учебная дисциплина дополняет знания, умения и навыки, получаемые по одновременно изучаемым и последующим дисциплинам:

- Анализ и оптимизация бизнес-процессов;
- Разработка и реализация проектов 2.

3 Планируемые результаты обучения по учебной дисциплине

Процесс изучения учебной дисциплины направлен на формирование следующих компетенций:

– Профессиональные компетенции

| Наименование категории (группы) ПК | Код и наименование ПК | Код и наименование индикатора достижения ПК | Планируемые результаты обучения |
|------------------------------------|---|---|---|
| | ПК-2: Способен разрабатывать и применять алгоритмы интеллектуального анализа больших объемов данных для | ПК-2.2 Собирает данные из различных источников и осуществляет их подготовку для анализа | – знать: основные методы выделения признаков, этапы проектирования признаков; |

| | | | |
|--|---------------------------------------|---|---|
| | управления технологическими системами | | <p> типовые методы моделирования и прогнозирования.</p> <p>– уметь: формировать признаки для решаемых задач; подбирать наиболее подходящий метод моделирования и прогнозирования.</p> <p>– владеть: навыками анализа предметной области, методами выделения признаков для моделей; типовыми методами моделирования и прогнозирования.</p> |
| | | ПК-2.3 Выбирает и использует методы и инструментальные средства для проведения анализа больших объемов данных, строит модели на основе данных | <p>– знать: особенности документирования данных для ML: признаков, методов, результатов моделирования и оценки.</p> <p>– уметь: документировать процесс разработки ML-модели.</p> <p>– владеть: навыками документирования процесса разработки ML-модели.</p> |

4 Объем и содержание учебной дисциплины

Учебные занятия по учебной дисциплине проводятся в форме контактной работы и в форме самостоятельной работы обучающихся.

Контактная работа обучающихся с педагогическим работником включает в себя занятия лекционного типа (лекции), занятия семинарского типа (семинары, практические занятия, практикумы), промежуточную аттестацию обучающихся и иную контактную работу, предусматривающую групповую или индивидуальную работу обучающихся с педагогическим работником. Контактная работа обучающихся с педагогическим работником может быть аудиторной, внеаудиторной, а также проводиться в электронной информационно-образовательной среде.

Объем учебной дисциплины

| Семестр / курс | | ИТОГО | 2 семестр |
|---|------------------------|--------------|--------------------------------|
| Форма промежуточной аттестации | | | экзамен, зачет с оценкой по КР |
| Трудоёмкость | <i>академ. час.</i> | 216 | 216 |
| | <i>зачетных единиц</i> | 6 | 6 |
| Лекции, <i>академ. час.</i> | | 16 | 16 |
| в форме практической подготовки | | 0 | 0 |
| Лабораторные работы, <i>академ. час.</i> | | 0 | 0 |
| в форме практической подготовки | | 0 | 0 |
| Практические занятия, <i>академ. час.</i> | | 48 | 48 |
| в форме практической подготовки | | 0 | 0 |
| Курсовая работа, <i>академ. час.</i> | | 36 | 36 |
| в форме практической подготовки | | 0 | 0 |
| Консультации, <i>академ. час.</i> | | 0 | 0 |
| в форме практической подготовки | | 0 | 0 |
| Самостоятельная работа, <i>академ. час.</i> | | 80 | 80 |
| в форме практической подготовки | | 0 | 0 |
| Контроль, <i>академ. час.</i> | | 36 | 36 |
| в форме практической подготовки | | 0 | 0 |

Содержание учебной дисциплины

Раздел 1 Сбор и подготовка реальных данных для машинного обучения;

Тема 1.1 Основные понятия машинного обучения (Основные термины и понятия. Подходы к обучению. Обзор средств и ресурсов для машинного обучения);

Тема 1.2 Сбор данных (Определение набора входных признаков. Наблюдаемое значение целевой переменной. Достаточный объем обучающих данных. Репрезентативность обучающей выборки);

Тема 1.3 Подготовка данных к моделированию (Категориальные признаки. Отсутствующие данные. Основы проектирования признаков. Нормализация данных);

Тема 1.4 Визуализация данных (Мозаичные диаграммы. Диаграммы размаха. Графики плотности. Диаграммы рассеяния);

Раздел 2 Моделирование и прогнозирование;

Тема 2.1 Основы моделирования с машинным обучением (Поиск связи между входными данными и целевой переменной. Типы методов моделирования. Обучение с учителем и без);

Тема 2.2 Классификация (Построение классификатора и получение предсказаний. Классификация сложных нелинейных данных. Классификация в случае множества классов);

Тема 2.3 Регрессия (Построение регрессора и генерация прогнозов. Регрессия для сложных нелинейных данных);

Раздел 3 Оценка и оптимизация модели;

Тема 3.1 Оценка прогностической точности на новых данных (Проблема переобучения и чрезмерно оптимистической оценки модели. Скользящий контроль. Перекрестная проверка);

Тема 3.2 Оценка моделей классификации (Точность для отдельных классов и таблица сопряженности. Компромиссы при оценке точности и ROC-кривые. Многоклассовая классификация);

Тема 3.3 Оценка моделей регрессии (Простые показатели эффективности регрессионных моделей. Исследование остатков);

Тема 3.4 Оптимизация модели путем подбора параметров (Параметры настройки ML-алгоритмов. Сеточный поиск);

Раздел 4 Основы проектирования признаков;

Тема 4.1 Основные этапы проектирования признаков (Проектирование признаков и знание предметной области. Этапы проектирования признаков. Извлечение признаков из обычного текста);

Тема 4.2 Выбор признаков (Прямой отбор и обратное исключение. Отбор признаков для исследования данных. Практический пример отбора признаков).

5 Перечень тем лекций

| № раздела / темы дисциплины | Темы лекций | Трудоемкость, <i>академ. час</i> | |
|-----------------------------|--|----------------------------------|---------------------------------|
| | | всего | в форме практической подготовки |
| Раздел 1. | Сбор и подготовка реальных данных для машинного обучения | | |
| Тема 1.1. | Основные понятия машинного обучения | 1 | |
| Тема 1.2. | Сбор данных | 1 | |
| Тема 1.3. | Подготовка данных к моделированию | 1 | |
| Тема 1.4. | Визуализация данных | 2 | |
| Раздел 2. | Моделирование и прогнозирование | | |
| Тема 2.1. | Основы моделирования с машинным обучением | 2 | |
| Тема 2.2. | Классификация | 1 | |

| | | | |
|---------------|---|-----------|----------|
| Тема 2.3. | Регрессия | 1 | |
| Раздел 3. | Оценка и оптимизация модели | | |
| Тема 3.1. | Оценка прогностической точности на новых данных | 1 | |
| Тема 3.2. | Оценка моделей классификации | 1 | |
| Тема 3.3. | Оценка моделей регрессии | 1 | |
| Тема 3.4. | Оптимизация модели путем подбора параметров | 2 | |
| Раздел 4. | Основы проектирования признаков | | |
| Тема 4.1. | Основные этапы проектирования признаков | 1 | |
| Тема 4.2. | Выбор признаков | 1 | |
| Итого: | | 16 | 0 |

6 Перечень тем практических занятий (семинаров)

| № раздела / темы дисциплины | Темы практических занятий (семинаров) | Трудоемкость, <i>академ. час</i> | |
|-----------------------------|---|----------------------------------|---------------------------------|
| | | всего | в форме практической подготовки |
| Тема 1.2. | Сбор и подготовка данных | 2 | |
| Тема 1.3. | Работа с различными типами данных в датасете | 2 | |
| Тема 1.4. | Типовые диаграммы для визуализации данных | 2 | |
| Тема 2.1. | Поиск связи между входными данными и целевой переменной | 4 | |
| Тема 2.2. | Построение классификатора и получение предсказаний | 6 | |
| Тема 2.3. | Построение регрессора и генерация прогнозов | 6 | |
| Тема 3.1. | Оценка прогностической точности на новых данных | 4 | |
| Тема 3.2. | Оценка моделей классификации | 4 | |
| Тема 3.3. | Оценка моделей регрессии | 4 | |
| Тема 3.4. | Оптимизация модели путем подбора параметров | 6 | |

| | | | |
|---------------|--|-----------|----------|
| Тема 4.1. | Этапы проектирования при-знаков | 4 | |
| Тема 4.2. | Прямой отбор и обратное исключение признаков для набора данных | 4 | |
| Итого: | | 48 | 0 |

7 Перечень тем лабораторных работ

| № раздела / темы дисциплины | Темы лабораторных работ | Трудоемкость, <i>академ. час</i> | |
|-----------------------------|-------------------------|----------------------------------|---------------------------------|
| | | всего | в форме практической подготовки |
| | <i>Отсутствуют</i> | | |
| Итого: | | 0 | 0 |

8 Перечень тем курсовых работ (проектов)

| № раздела / темы дисциплины | Темы курсовых работ (проектов) | Трудоемкость, <i>академ. час</i> | |
|-------------------------------------|---|----------------------------------|---------------------------------|
| | | всего | в форме практической подготовки |
| Раздел 1; Раздел 2; Раздел 3. | 1. Разработка системы прогноза инсульта у пациента 2. Разработка рекомендательной системы подбора электроники 3. Разработка системы сегментации клиентов 4. Разработка рекомендательной системы ответов на профессиональные вопросы 5. Разработка системы прогнозирования пригодности для использования питьевой воды | 36 | |
| Итого: | | 36 | 0 |

9 Виды самостоятельной работы

| № раздела / темы дисциплины | Виды самостоятельной работы | Трудоемкость, <i>академ. час</i> | |
|-----------------------------|-----------------------------|----------------------------------|---------------------------------|
| | | всего | в форме практической подготовки |
| Раздел 1. | 1. Изучение теоретического | 18 | |

| | | | |
|------------------------|---|------------|----------|
| | материала; 2. Подготовка к практическому занятию; 3. Прохождение тестирования. | | |
| Раздел 2. | 1. Изучение теоретического материала; 2. Подготовка к практическому занятию; 3. Прохождение тестирования. | 22 | |
| Раздел 3. | 1. Изучение теоретического материала; 2. Подготовка к практическому занятию; 3. Прохождение тестирования. | 22 | |
| Раздел 4. | 1. Изучение теоретического материала; 2. Подготовка к практическому занятию; 3. Прохождение тестирования. | 18 | |
| <i>Курсовая работа</i> | <i>Выполнение курсовой работы</i> | 36 | 0 |
| <i>Контроль</i> | <i>Подготовка к экзамену</i> | 36 | |
| Итого: | | 152 | 0 |

10 Учебно-методическое и информационное обеспечение учебной дисциплины

а) литература:

1 Флах, П. Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных : учебное пособие / П. Флах. – Москва : ДМК-пресс, 2015. – 400 с. – ISBN 978-5-97060-273-7. – URL: <https://www.studentlibrary.ru/book/ISBN9785970602737.html> (дата обращения: 09.03.2022);

2 Рашка, С. Python и машинное обучение: крайне необходимое пособие по новейшей предсказательной аналитике, обязательное для более глубокого понимания методологии машинного обучения : практическое пособие / С. Рашка. – Москва : ДМК-пресс, 2017. – 418 с. – ISBN 978-5-97060-409-0. – URL: <https://www.studentlibrary.ru/book/ISBN9785970604090.html> (дата обращения: 09.03.2022);

3 Паттерсон, Дж. Глубокое обучение с точки зрения практика / Дж. Паттерсон, А. Гибсон. – Москва : ДМК-пресс, 2018. – ISBN 978-5-97060-

481-6. – URL: <https://www.studentlibrary.ru/book/ISBN9785970604816.html> (дата обращения: 09.03.2022);

4 Хейдт, М. Изучаем pandas / М. Хейдт, А. В. Груздев. – Москва : ДМК-пресс, 2019. – ISBN 978-5-97060-670-4. – URL: <https://www.studentlibrary.ru/book/ISBN9785970606704.html> (дата обращения: 09.03.2022).

б) ресурсы информационно-телекоммуникационной сети «Интернет»:

1 Консультант студента. Электронная библиотека технического ВУЗа : электронно-библиотечная система / ООО «Политехресурс». – Москва, [200 –]. – URL: <http://www.studentlibrary.ru>. – Режим доступа: для авторизир. пользователей;

2 ЛАНЬ : электронно-библиотечная система : [коллекция «Инженерно-технические науки»] / ООО «Издательство Лань». – Санкт-Петербург, [200 –]. – URL: <http://e.lanbook.com>. – Режим доступа: для авторизир. пользователей;

3 НАУЧНАЯ ЭЛЕКТРОННАЯ БИБЛИОТЕКА eLIBRARY.RU : база данных / ООО «НЭБ». – Москва, [200 –]. – URL: <http://elibrary.ru>. – Режим доступа: по подписке;

4 Образовательная платформа ЮРАЙТ / ООО «Электронное издательство Юрайт». – Москва, [200 –]. – URL: <https://urait.ru>. – Режим доступа: для авторизир. пользователей;

5 Университетская библиотека онлайн : электронно-библиотечная система / ООО «Директ-Медиа». – Москва, [200 –]. – URL: <http://www.biblioclub.ru>. – Режим доступа: для авторизир. пользователей;

6 Электронная библиотека // Научно-техническая библиотека СибГИУ : сайт. – Новокузнецк, [200 –]. – URL: <http://library.sibsiu.ru/LibrELibraryFullText.asp>. – Режим доступа: для авторизир. пользователей;

7 Электронный каталог : сайт / Научно-техническая библиотека СибГИУ. – Новокузнецк, [199 –]. – URL: <http://libr.sibsiu.ru>.

в) лицензионное и свободно распространяемое программное обеспечение:

- 7-Zip;
- Adobe Acrobat Reader;
- Kaspersky Endpoint Security;
- Microsoft Office 2007;
- Microsoft Windows 7;
- Notepad++.

г) базы данных и информационно-справочные системы:

1 КонсультантПлюс : справочно-правовая система / ООО «Информационный центр АНВИК». – Новокузнецк, [199 –]. – Режим доступа: компьютерная сеть библиотеки Сиб. гос. индустр. ун-та.;

2 Система ГАРАНТ : электронный периодический справочник / ООО «Правовой центр «Гарант». – Кемерово, [200 –]. – Режим доступа: компьютерная сеть Сиб. гос. индустр. ун-та.;

3 Техэксперт : информационно-справочная система / ООО «Группа компаний «Кодекс». – Кемерово, [200 –]. – Режим доступа: компьютерная сеть Сиб. гос. индустр. ун-та.;

4 Электронный реферативный журнал (ЭлРЖ) : база данных / ВИНТИ РАН. – Москва, [200 –]. – Режим доступа: компьютерная сеть библиотеки Сиб. гос. индустр. ун-та.

11 Материально-техническое обеспечение учебной дисциплины

Материально-техническое обеспечение учебной дисциплины включает учебные аудитории, оснащенные оборудованием, компьютерной техникой, и техническими средствами обучения, в том числе:

- учебную аудиторию для проведения занятий лекционного типа, оборудованную учебной доской, экраном и мультимедийным проектором;
- учебную аудиторию для проведения занятий семинарского типа (практических занятий), оснащенную персональными компьютерами и выходом в сеть «Интернет»;
- учебную аудиторию для выполнения курсовых работ;
- учебную аудиторию (помещения) для групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации;
- помещения для самостоятельной работы, оснащенные компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду, научно-техническую библиотеку СибГИУ.

Рабочая программа дисциплины составлена в соответствии с требованиями ФГОС ВО по направлению подготовки (специальности) 09.04.03 «Прикладная информатика».

Составитель(и):

доцент Бабичева Надежда Борисовна (кафедра прикладных информационных технологий и программирования).

Рабочая программа дисциплины рассмотрена и утверждена на заседании кафедры.

Приложение А

Аннотация рабочей программы дисциплины «Машинное обучение для задач управления»

по направлению подготовки (специальности)
09.04.03 «Прикладная информатика»
(направленность (профиль): «Системы корпоративного
управления»)
форма обучения – Очная форма

1 Цели и задачи освоения учебной дисциплины

Целями учебной дисциплины являются:

- формирование у обучающихся знаний и навыков применения методов машинного обучения для задач управления; моделирования и прогнозирования ML-методами, а также оценки построенной модели.

Задачами учебной дисциплины являются:

- изучение методов сбора и подготовки реальных данных для машинного обучения;
- изучение этапов проектирования признаков для построения моделей машинного обучения;
- изучение процесса построения моделей классификации и регрессии методами ML;
- изучение подходов к оценке построенных моделей.

2 Место учебной дисциплины в структуре ООП по направлению подготовки (специальности)

Учебная дисциплина относится к учебным дисциплинам части, формируемой участниками образовательных отношений **Блока 1 «Дисциплины (модули)»** ООП по направлению подготовки (специальности) 09.04.03 «Прикладная информатика».

Учебная дисциплина базируется на предварительном усвоении обучающимися учебных дисциплин:

- Python для анализа данных;
- Методология научного познания.

Учебная дисциплина дополняет знания, умения и навыки, получаемые по одновременно изучаемым и последующим дисциплинам:

- Анализ и оптимизация бизнес-процессов;
- Разработка и реализация проектов 2.

3 Планируемые результаты обучения по учебной дисциплине

Процесс изучения учебной дисциплины направлен на формирование следующих компетенций:

– Профессиональные компетенции

| Наименование категории (группы) ПК | Код и наименование ПК | Код и наименование индикатора достижения ПК | Планируемые результаты обучения |
|------------------------------------|---|--|--|
| | ПК-2: Способен разрабатывать и применять алгоритмы интеллектуального анализа больших объемов данных для управления технологическими системами | ПК-2.2 Собирает данные из различных источников и осуществляет их подготовку для анализа | <p>– знать: основные методы выделения признаков, этапы проектирования признаков; типовые методы моделирования и прогнозирования.</p> <p>– уметь: формировать признаки для решаемых задач; подбирать наиболее подходящий метод моделирования и прогнозирования.</p> <p>– владеть: навыками анализа предметной области, методами выделения признаков для моделей; типовыми методами моделирования и прогнозирования.</p> |
| | ПК-2.3 Выбирает и использует методы и инструментальные средства для проведения анализа больших объемов данных, строит модели на основе данных | <p>– знать: особенности документирования данных для ML: признаков, методов, результатов моделирования и оценки.</p> <p>– уметь: документировать процесс разработки ML-модели.</p> <p>– владеть: навыками</p> | |

| | | | |
|--|--|--|---|
| | | | документирования процесса разработки ML-модели. |
|--|--|--|---|

4 Объем учебной дисциплины

| Семестр / курс | | ИТОГО | 2 семестр |
|---|------------------------|--------------|--------------------------------|
| Форма промежуточной аттестации | | | экзамен, зачет с оценкой по КР |
| Трудоёмкость | <i>академ. час.</i> | 216 | 216 |
| | <i>зачетных единиц</i> | 6 | 6 |
| Лекции, <i>академ. час.</i> | | 16 | 16 |
| в форме практической подготовки | | 0 | 0 |
| Лабораторные работы, <i>академ. час.</i> | | 0 | 0 |
| в форме практической подготовки | | 0 | 0 |
| Практические занятия, <i>академ. час.</i> | | 48 | 48 |
| в форме практической подготовки | | 0 | 0 |
| Курсовая работа, <i>академ. час.</i> | | 36 | 36 |
| в форме практической подготовки | | 0 | 0 |
| Консультации, <i>академ. час.</i> | | 0 | 0 |
| в форме практической подготовки | | 0 | 0 |
| Самостоятельная работа, <i>академ. час.</i> | | 80 | 80 |
| в форме практической подготовки | | 0 | 0 |
| Контроль, <i>академ. час.</i> | | 36 | 36 |
| в форме практической подготовки | | 0 | 0 |

5 Краткое содержание учебной дисциплины

В структуре учебной дисциплины выделяются следующие основные разделы (темы):

Раздел 1 Сбор и подготовка реальных данных для машинного обучения;

Тема 1.1 Основные понятия машинного обучения (Основные термины и понятия. Подходы к обучению. Обзор средств и ресурсов для машинного обучения);

Тема 1.2 Сбор данных (Определение набора входных признаков. Наблюдаемое значение целевой переменной. Достаточный объем обучающих данных. Репрезентативность обучающей выборки);

Тема 1.3 Подготовка данных к моделированию (Категориальные признаки. Отсутствующие данные. Основы проектирования признаков. Нормализация данных);

Тема 1.4 Визуализация данных (Мозаичные диаграммы. Диаграммы размаха. Графики плотности. Диаграммы рассеяния);

Раздел 2 Моделирование и прогнозирование;

Тема 2.1 Основы моделирования с машинным обучением (Поиск связи между входными данными и целевой переменной. Типы методов моделирования. Обучение с учителем и без);

Тема 2.2 Классификация (Построение классификатора и получение предсказаний. Классификация сложных нелинейных данных. Классификация в случае множества классов);

Тема 2.3 Регрессия (Построение регрессора и генерация прогнозов. Регрессия для сложных нелинейных данных);

Раздел 3 Оценка и оптимизация модели;

Тема 3.1 Оценка прогностической точности на новых данных (Проблема переобучения и чрезмерно оптимистической оценки модели. Скользящий контроль. Перекрестная проверка);

Тема 3.2 Оценка моделей классификации (Точность для отдельных классов и таблица сопряженности. Компромиссы при оценке точности и ROC-кривые. Многоклассовая классификация);

Тема 3.3 Оценка моделей регрессии (Простые показатели эффективности регрессионных моделей. Исследование остатков);

Тема 3.4 Оптимизация модели путем подбора параметров (Параметры настройки ML-алгоритмов. Сеточный поиск);

Раздел 4 Основы проектирования признаков;

Тема 4.1 Основные этапы проектирования признаков (Проектирование признаков и знание предметной области. Этапы проектирования признаков. Извлечение признаков из обычного текста);

Тема 4.2 Выбор признаков (Прямой отбор и обратное исключение. Отбор признаков для исследования данных. Практический пример отбора признаков).

6 Составитель(и):

доцент Бабичева Надежда Борисовна (кафедра прикладных информационных технологий и программирования).