|  |
| --- |
| утверждаю  Руководитель Департамента  анализа данных, принятия решений и финансовых технологий  \_\_\_\_\_\_\_\_\_\_\_\_ В.И. Соловьев  15.10.2019 г. |
|  |

ПРИЛОЖЕНИЕ К РАБОЧЕЙ ПРОГРАММЕ ДИСЦИПЛИНЫ

**ТЕХНОЛОГИИ АНАЛИЗА ДАННЫХ**

**И МАШИННОГО ОБУЧЕНИЯ**

НАПРАВЛЕНИЕ ПОДГОТОВКИ: **09.03.03 «Прикладная информатика»**

ПРОФИЛИ: **«****Высокопроизводительные вычисления в цифровой экономике»,** «**ИТ-сервисы и технологии обработки данных в экономике и финансах»**

ФОРМА ОБУЧЕНИЯ: **Очная, очно-заочная, заочная**

ГОД ПРИЕМА: **2017, 2018, 2019, 2020**

Год утверждения программы: 2017 год

*Одобрено департаментом анализа данных, принятия решений и финансовых технологий*

*Протокол № 3 от 15.10.2019 г*

**Содержание**

|  |  |
| --- | --- |
| **Наименование разделов РПД** | **стр.** |
| Перечень планируемых результатов освоения образовательной программы (перечень компетенций) с указанием индикаторов их достижения и планируемых результатов обучения по дисциплине | 2 |
| Объем дисциплины в зачетных единицах и в академических часах с выделением объема аудиторной (лекции, семинары) и самостоятельной работы обучающихся (в семестре, в сессию) | 5 |
| Учебно-тематический план | 7 |
| Содержание семинаров, практических занятий | 13 |
| Перечень вопросов, отводимых на самостоятельное освоение дисциплины, формы внеаудиторной самостоятельной работы | 15 |
| Перечень вопросов, заданий, тем для подготовки к текущему контролю | 17 |
| Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине | 18 |
| Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень необходимого программного обеспечения и информационных справочных систем | 24 |

**2. Перечень планируемых результатов освоения образовательной программы с указанием индикаторов их достижения и планируемых результатов обучения по дисциплине**

***Профиль «ИТ-сервисы и технологии обработки данных в экономике и финансах»***

|  |  |  |  |
| --- | --- | --- | --- |
| **Код компе**  **тенции** | **Наименование**  **компетенции** | **Индикаторы достижения компетенции** | **Результаты обучения (владения, умения и знания), соотнесенные с компетенциями/ индикаторами достижения компетенции** |
| **2017, 2018 год приема** | | | |
| ПК-7 | Способность проводить описание прикладных процессов и информационного обеспечения решения прикладных задач |  | **Знать**: процессы собора исходных данных о задаче из области экономики и финансов для создания и внедрения модели МО  **Уметь**: выполнять процессом собора исходных данных о задаче из области экономики и финансов для создания и внедрения модели МО  **Владеть**: современными подходами к процессу собора исходных данных о задаче из области экономики и финансов для создания и внедрения модели МО |
| ПК-8 | Способность программировать приложения и создавать программные прототипы решения прикладных задач | - | **Знать:** различные современные алгоритмы машинного обучения.  **Уметь:** выбирать наиболее адекватные алгоритмы машинного обучения.  **Владеть:** современными технологиями построения алгоритмов машинного обучения. |
| ПК-14 | Способность осуществлять ведение базы данных и поддержку информационного обеспечения решения прикладных задач | - | **Знать**: Логику ведения базы данных  **Уметь**: Осуществлять ведение базы данных  **Владеть**: способностью выполнять поддержку информационного обеспечения решения прикладных задач с помощью современных СУБД |
| ПК-23 | Способность применять системный подход и математические методы в формализации решения прикладных задач | - | **Знать**: системный подход для решения прикладных задач  **Уметь**: применять математические методы в решении прикладных задач  **Владеть**: подходами к формализации решения прикладных задач |
| ПКП-2 | Способность применять методы внедрения и эксплуатации корпоративных информационных систем в сфере экономики и финансов | - | **Знать**: процессы определения качества модели МО  **Уметь**: выполнять процессом определения качества модели МО  **Владеть**: современными подходами к процессу определения качества модели МО |
| ПКП-4 | Способность применять технологии моделирования и анализа процессов в сфере экономики и финансов | - | **Знать**: технологии моделирования и анализа процессов  **Уметь**: применять технологии моделирования и анализа процессов  **Владеть**: использованием технологий моделирования и анализа процессов в сфере экономики и финансов |
| **2019, 2020 год приема** | | | |
| ПКП-2 | Способность применять методы внедрения и эксплуатации корпоративных информационных систем в сфере экономики и финансов | 1.Демонстрирует знание назначения и функционал типовых модулей корпоративных информационных систем, основные этапы и методы их внедрения и эксплуатации  2.Владеет методологией внедрения и эксплуатации приложений в сфере экономики и финансов на платформе корпоративных информационных систем. | **Знать:** назначения и функционал типовых модулей корпоративных информационных систем  **Уметь:** выполнять основные этапы и методы внедрения и эксплуатации типовых модулей корпоративных информационных систем  **Знать:** методологию внедрения и эксплуатации приложений  **Уметь:** внедрять и эксплуатировать приложения в сфере экономики и финансов на платформе корпоративных информационных систем |
| ПКП-4 | Способность применять технологии моделирования и анализа процессов в сфере экономики и финансов | 1.Демонстрирует знания в области теории и методологии моделирования и анализа в сфере экономики и финансов.  2.Строит математические модели в сфере экономики и финансов | **Знать:** теорию и методологии моделирования и анализа  **Уметь:** выполнять моделирование и анализ в сфере экономики и финансов  **Знать:** математические модели  **Уметь**: строить математические модели в сфере экономики и финансов |

***Профиль «Высокопроизводительные вычисления в цифровой экономике»***

|  |  |  |  |
| --- | --- | --- | --- |
| **Код компе**  **тенции** | **Наименование**  **компетенции** | **Индикаторы достижения компетенции** | **Результаты обучения (владения, умения и знания), соотнесенные с компетенциями/ индикаторами достижения компетенции** |
| ***2018 год приема*** | | | |
| ПК-14 | Способность осуществлять ведение базы данных и поддержку информационного обеспечения решения прикладных задач | **-** | **Знать**: Логику ведения базы данных  **Уметь**: Осуществлять ведение базы данных  **Владеть**: способностью выполнять поддержку информационного обеспечения решения прикладных задач с помощью современных СУБД |
| ПК-20 | Способность осуществлять и обосновывать выбор проектных решений по видам обеспечения информационных систем | **-** | **Знать**: проектные решения по видам обеспечения информационных систем  **Уметь**: осуществлять и обосновывать выбор проектных решений  **Владеть**: Способностью обосновывать выбор проектных решений по видам обеспечения информационных систем |
| ПКП-7 | Способность провести статистический анализ данных, в том числе разнородных и данных большого объема | **-** | **Знать**: методы статистического анализа данных  **Уметь**: проводить статистический анализ данных, в том числе разнородных  **Владеть**: Способностью проводить статистический анализ данных, в том числе большого объема данных |
| ***2019, 2020 год приема*** | | | |
| ОПК-4 | Способен участвовать в разработке стандартов, норм и правил, а также технической документации, связанной с профессиональной деятельностью | 1.Разрабатывает техническую документацию для программных компонентов и информационных систем.  2. Работает со стандартами, в том числе адаптирует стандарты для специфических требований разрабатываемого программного обеспечения. | **Знать: методологию описания** программных компонентов и информационных систем  **Уметь:** разрабатывать техническую документацию  **Знать:** стандарты для специфических требований разрабатываемого программного обеспечения  **Уметь:** работать со стандартами, в том числе адаптировать их |
| ПКП-7 | Способность провести статистический анализ данных, в том числе разнородных и данных большого объема | 1. Демонстрирует методы анализа статистических данных, рассчитывает показатели качества статистического анализа данных в зависимости от объема выборки.  2. Разрабатывает алгоритмы для обработки больших данных.  3. Обосновывает выбор методов статистического анализа в зависимости от типов данных, в том числе нечисловых данных. | **Знать:** методы анализа статистических данных  **Уметь:** рассчитывать показатели качества статистического анализа данных в зависимости от объема выборки  **Знать:** алгоритмы для обработки больших данных  **Уметь:** разрабатывать алгоритмы для обработки больших данных  **Знать:** методы статистического анализа данных  **Уметь:** проводить статистический анализ данных, в том числе разнородны хи большого объема |

**4. Объем дисциплины в зачетных единицах и в академических часах с выделением объема аудиторной (лекции, семинары) и самостоятельной работы обучающихся (в семестре, в сессию)**

***Профиль «ИТ-сервисы и технологии обработки данных в экономике и финансах»***

Общая трудоемкость дисциплины составляет 7 зачетных единиц.

Вид промежуточной аттестации: экзамены (5, 6 семестр)

Вид текущего контроля – контрольная работа

*Очная форма обучения, 2017 г.*

|  |  |  |  |
| --- | --- | --- | --- |
| **Вид учебной работы по дисциплине** | **Всего**  **(в з/е и часах)** | **Семестр 5**  **(в часах)** | **Семестр 6**  **(в часах)** |
| **Общая трудоемкость дисциплины** | **7/252** | **126** | **126** |
| ***Контактная работа - Аудиторные занятия*** | ***108*** | ***54*** | ***54*** |
| *Лекции* | *36* | *18* | *18* |
| *Семинары, практические занятия* | *72* | *36* | *36* |
| ***Самостоятельная работа*** | ***144*** | ***72*** | ***72*** |
| Вид текущего контроля | Контрольная  работа | Контрольная  работа | *-* |
| Вид промежуточной аттестации | Экзамен | Экзамен | Экзамен |

Общая трудоемкость дисциплины составляет 7 зачетных единиц.

Вид промежуточной аттестации: экзамены (5, 6 семестр)

Вид текущего контроля – контрольные работы (5, 6 семестр)

*Очная форма обучения 2018 г.*

|  |  |  |  |
| --- | --- | --- | --- |
| **Вид учебной работы по дисциплине** | **Всего**  **(в з/е и часах)** | **Семестр 5**  **(в часах)** | **Семестр 6**  **(в часах)** |
| **Общая трудоемкость дисциплины** | **7/252** | **126** | **126** |
| ***Контактная работа - Аудиторные занятия*** | ***100*** | ***50*** | ***50*** |
| *Лекции* | *32* | *16* | *16* |
| *Семинары, практические занятия* | *68* | *34* | *34* |
| ***Самостоятельная работа*** | ***152*** | ***76*** | ***76*** |
| Вид текущего контроля | Контрольные  работы | Контрольная  работа | Контрольная  работа |
| Вид промежуточной аттестации | Экзамен | Экзамен | Экзамен |

Общая трудоемкость дисциплины составляет 6 зачетных единиц.

Вид промежуточной аттестации: экзамен (6, 7 семестр)

Вид текущего контроля – курсовая работа, контрольная работа

*Очная / очно-заочная форма обучения 2019 г.*

|  |  |  |  |
| --- | --- | --- | --- |
| **Вид учебной работы по дисциплине** | **Всего**  **(в з/е и часах)** | **Семестр 6**  **(в часах)** | **Семестр 7**  **(в часах)** |
| **Общая трудоемкость дисциплины** | **6/216** | **108** | **108** |
| ***Контактная работа –***  ***Аудиторные занятия*** | ***100/68*** | ***50/34*** | ***50/34*** |
| *Лекции* | *32/32* | *16/16* | *16/16* |
| *Семинары, практические занятия* | *68//36* | *34/18* | *34/18* |
| ***Самостоятельная работа*** | ***116/148*** | ***58/74*** | ***58/74*** |
| Вид текущего контроля | Курсовая работа, контрольная работа | Курсовая работа  (20 час.) | Контрольная  работа |
| Вид промежуточной аттестации | Экзамен | Экзамен | Экзамен |

Общая трудоемкость дисциплины составляет 8 зачетных единиц.

Вид промежуточной аттестации: экзамены (5, 6 семестр)

Вид текущего контроля – курсовая работа, контрольная работа

*Очная / очно-заочная форма обучения 2020 г.*

|  |  |  |  |
| --- | --- | --- | --- |
| **Вид учебной работы по дисциплине** | **Всего**  **(в з/е и часах)** | **Семестр 5**  **(в часах)** | **Семестр 6**  **(в часах)** |
| **Общая трудоемкость дисциплины** | **8/288** | **144** | **144** |
| ***Контактная работа –***  ***Аудиторные занятия*** | ***100/66*** | ***50/34*** | ***50/32*** |
| *Лекции* | *32/32* | *16/16* | *16/16* |
| *Семинары, практические занятия* | *68/34* | *34/18* | *34/16* |
| ***Самостоятельная работа*** | ***188/222*** | ***94/110*** | ***94/112*** |
| Вид текущего контроля | Контрольная работа, курсовая работа | *Контрольная работа* | *Курсовая работа*  *(20 час.)* |
| Вид промежуточной аттестации | Экзамен | Экзамен | Экзамен |

**Профиль «Высокопроизводительные вычисления в цифровой экономике»**

Общая трудоемкость дисциплины составляет 5 зачетных единиц.

Вид промежуточной аттестации: экзамен

Вид текущего контроля – расчетно-аналитическая работа

|  |  |  |
| --- | --- | --- |
| **Вид учебной работы по дисциплине** | **Всего**  **(в з/е и часах)** | **Семестр 8**  **(в часах)** |
| **Общая трудоемкость дисциплины** | **5/180** | **180** |
| ***Контактная работа –***  ***Аудиторные занятия*** | ***20*** | ***20*** |
| *Лекции* | *4* | *4* |
| *Семинары, практические занятия* | *16* | *16* |
| ***Самостоятельная работа*** | ***160*** | ***160*** |
| Вид текущего контроля | Курсовая работа  (20 час) | Курсовая работа  (20 час) |
| Вид промежуточной аттестации | Экзамен | Экзамен |

*Заочная форма обучения 2018, 2019, 2020 г.*

**5.2.** **Учебно – тематический план**

***Профиль «ИТ-сервисы и технологии обработки данных в экономике и финансах»***

*Очная форма обучения, 2017 г.*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **№**  **п/п** | **Наименование тем (разделов)**  **дисциплины** | **Трудоёмкость в часах** | | | | | | **Формы текущего контроля успеваемости** |
| **Всего** | **Аудиторная работа** | | | | **Самостоятельная работа** |
| Общая, в т. ч.: | Лекции | Семинары, практические занятия | Занятия в интерактивной форме |
| 1. | Информационные технологии анализа данных | 26 | 12 | 4 | 8 | 4 | 14 | Устный опрос, проверка практических заданий |
| 2. | Введение в машинное обучение и знакомство с технологиями машинного обучения | 20 | 6 | 2 | 4 | 4 | 14 | Устный опрос, проверка практических заданий |
| 3. | Методы обучения с учителем | 26 | 12 | 4 | 8 | 8 | 14 | Устный опрос, проверка практических заданий |
| 4. | Методы восстановления регрессии | 26 | 12 | 4 | 8 | 4 | 14 | Устный опрос, проверка практических заданий |
| 5. | Обучение без учителя, кластеризация | 28 | 12 | 4 | 8 | 4 | 16 | Устный опрос, проверка практических заданий |
| 6. | Коллаборативная фильтрация, отбор признаков | 20 | 6 | 2 | 4 | 4 | 14 | Устный опрос, проверка практических заданий |
| 7. | Генетические алгоритмы | 20 | 6 | 2 | 4 | 4 | 14 | Устный опрос, проверка практических заданий |
| 8. | Анализ сетевых структуру | 32 | 18 | 6 | 12 | 8 | 14 | Устный опрос, проверка практических заданий |
| 9. | Анализ текстов на естественном языке | 26 | 12 | 4 | 8 | 4 | 14 | Устный опрос, проверка практических заданий |
| 10. | Визуализация результатов анализа данных | 28 | 12 | 4 | 8 | 4 | 16 | Устный опрос, проверка практических заданий |
|  | В целом по дисциплине | 252 | 108 | 36 | 72 | 48 | 144 | Контрольная работа |
|  | Итого в % |  |  |  |  | 52% |  |  |

*Очная форма обучения, 2018 г.*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **№**  **п/п** | **Наименование тем (разделов)**  **дисциплины** | **Трудоёмкость в часах** | | | | | | **Формы текущего контроля успеваемости** |
| **Всего** | **Аудиторная работа** | | | | **Самостоятельная работа** |
| Общая, в т.ч. | Лекции | Семинары, практические занятия | Занятия в интерактивной форме |
| 1. | Информационные технологии анализа данных | 26 | 8 | 2 | 6 | 2 | 18 | Устный опрос, проверка практических заданий |
| 2. | Введение в машинное обучение и знакомство с технологиями машинного обучения | 20 | 6 | 2 | 4 | 2 | 14 | Устный опрос, проверка практических заданий |
| 3. | Методы обучения с учителем | 26 | 12 | 4 | 8 | 4 | 14 | Устный опрос, проверка практических заданий |
| 4. | Методы восстановления регрессии | 26 | 12 | 4 | 8 | 2 | 14 | Устный опрос, проверка практических заданий |
| 5. | Обучение без учителя, кластеризация | 28 | 12 | 4 | 8 | 2 | 16 | Устный опрос, проверка практических заданий |
| 6. | Коллаборативная фильтрация, отбор признаков | 20 | 6 | 2 | 4 | 2 | 14 | Устный опрос, проверка практических заданий |
| 7. | Генетические алгоритмы | 20 | 6 | 2 | 4 | 2 | 14 | Устный опрос, проверка практических заданий |
| 8. | Анализ сетевых структуру | 32 | 14 | 4 | 10 | 4 | 18 | Устный опрос, проверка практических заданий |
| 9. | Анализ текстов на естественном языке | 26 | 12 | 4 | 8 | 2 | 14 | Устный опрос, проверка практических заданий |
| 10. | Визуализация результатов анализа данных | 28 | 12 | 4 | 8 | 2 | 16 | Устный опрос, проверка практических заданий |
|  | В целом по дисциплине | 252 | 100 | 32 | 68 | 24 | 152 | Контрольные работы |
|  | Итого в % |  |  |  |  | 24% |  |  |

*Очная / очно-заочная форма обучения , 2019*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **№**  **п/п** | **Наименование тем**  **(разделов)**  **дисциплины** | **Трудоёмкость в часах** | | | | | | **Формы текущего контроля успеваемости** |
| **Всего** | **Аудиторная работа** | | | | **Самостоятельная работа** |
| Общая, в т.ч.: | Лекции | Семинары, практические занятия | Занятия в интерактивной форме |
| 1. | Информационные технологии анализа данных | 19/20 | 8/6 | 2 | 6/4 | 2 | 11/14 | Устный опрос, проверка практических заданий |
| 2. | Введение в машинное обучение и знакомство с технологиями машинного обучения | 19/20 | 8/6 | 2 | 6/4 | 2 | 11/14 | Устный опрос, проверка практических заданий |
| 3. | Методы обучения с учителем | 21/22 | 10/8 | 4 | 6/4 | 2 | 11/14 | Устный опрос, проверка практических заданий |
| 4. | Методы восстановления регрессии | 23/22 | 12/8 | 4 | 8/4 | 2 | 11/14 | Устный опрос, проверка практических заданий |
| 5. | Обучение без учителя, кластеризация | 26/24 | 12/6 | 4 | 8/2 | 2 | 14/18 | Устный опрос, проверка практических заданий |
| 6. | Коллаборативная фильтрация, отбор признаков | 19/20 | 8/6 | 2 | 6/4 | 2 | 11/14 | Устный опрос, проверка практических заданий |
| 7. | Генетические алгоритмы | 19/20 | 8/6 | 2 | 6/4 | 2 | 11/14 | Устный опрос, проверка практических заданий |
| 8. | Анализ сетевых структуру | 21/20 | 10/6 | 4 | 6/2 | 2 | 11/14 | Устный опрос, проверка практических заданий |
| 9. | Анализ текстов на естественном языке | 23/22 | 12/8 | 4 | 8/4 | 2 | 11/14 | Устный опрос, проверка практических заданий |
| 10. | Визуализация результатов анализа данных | 26/26 | 12/8 | 4 | 8/4 | 4 | 14/18 | Устный опрос, проверка практических заданий |
|  | В целом по дисциплине | 216 | 100/68 | 32 | 68/36 | 22 | 116/148 | Курсовая работа, контрольная работа |
|  | Итого в % |  |  |  |  | 22/32% |  |  |

*Очная / очно-заочная форма обучения, 2020 г.*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **№**  **п/п** | **Наименование тем**  **(разделов)**  **дисциплины** | **Трудоёмкость в часах** | | | | | | **Формы текущего контроля успеваемости** |
| **Всего** | **Аудиторная работа** | | | | **Самостоятельная работа** |
| Общая, в т.ч.: | Лекции | Семинары, практические занятия | Занятия в интерактивной форме |
| 1. | Информационные технологии анализа данных | 32/28 | 8/6 | 2 | 6/4 | 2 | 24/22 | Устный опрос, проверка практических заданий |
| 2. | Введение в машинное обучение и знакомство с технологиями машинного обучения | 28/26 | 8/6 | 2 | 6/4 | 2 | 20/20 | Устный опрос, проверка практических заданий |
| 3. | Методы обучения с учителем | 26/34 | 10/8 | 4 | 6/4 | 2 | 16/26 | Устный опрос, проверка практических заданий |
| 4. | Методы восстановления регрессии | 30/30 | 12/8 | 4 | 8/4 | 2 | 18/22 | Устный опрос, проверка практических заданий |
| 5. | Обучение без учителя, кластеризация | 30/28 | 12/6 | 4 | 8/2 | 2 | 18/22 | Устный опрос, проверка практических заданий |
| 6. | Коллаборативная фильтрация, отбор признаков | 26/28 | 8/6 | 2 | 6/4 | 2 | 18/22 | Устный опрос, проверка практических заданий |
| 7. | Генетические алгоритмы | 26/28 | 8/6 | 2 | 6/4 | 2 | 18/22 | Устный опрос, проверка практических заданий |
| 8. | Анализ сетевых структуру | 30/28 | 10/6 | 4 | 6/2 | 2 | 20/22 | Устный опрос, проверка практических заданий |
| 9. | Анализ текстов на естественном языке | 30/30 | 12/8 | 4 | 8/4 | 2 | 18/22 | Устный опрос, проверка практических заданий |
| 10. | Визуализация результатов анализа данных | 30/28 | 12/6 | 4 | 8/2 | 2 | 18/22 | Устный опрос, проверка практических заданий |
|  | В целом по дисциплине | 288 | 100/66 | 32 | 68/34 | 20 | 188/222 | Контрольная работа, курсовая работа |
|  | Итого в % |  |  |  |  | 20/30% |  |  |

**Профиль «Высокопроизводительные вычисления в цифровой экономике»**

*Заочная форма обучения 2018, 2019, 2020 г.*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **№**  **п/п** | **Наименование тем**  **(разделов)**  **дисциплины** | **Трудоёмкость в часах** | | | | | | **Формы текущего контроля успеваемости** |
| **Всего** | **Аудиторная работа** | | | | **Самостоятельная работа** |
| Общая, в т.ч.: | Лекции | Семинары, практические занятия | Занятия в интерактивной форме |
| 1. | Информационные технологии анализа данных | 10 | 2 | 2 | - | - | 8 | Устный опрос, проверка практических заданий |
| 2. | Введение в машинное обучение и знакомство с технологиями машинного обучения | 14 | 2 | 2 | - | - | 12 | Устный опрос, проверка практических заданий |
| 3. | Методы обучения с учителем | 18 | 2 | - | 2 | 1 | 16 | Устный опрос, проверка практических заданий |
| 4. | Методы восстановления регрессии | 11 | 1 | - | 1 | - | 10 | Устный опрос, проверка практических заданий |
| 5. | Обучение без учителя, кластеризация | 12 | 2 | - | 2 | 1 | 10 | Устный опрос, проверка практических заданий |
| 6. | Коллаборативная фильтрация, отбор признаков | 11 | 1 | - | 1 | - | 10 | Устный опрос, проверка практических заданий |
| 7. | Генетические алгоритмы | 15 | 1 | - | 1 | - | 14 | Устный опрос, проверка практических заданий |
| 8. | Анализ сетевых структуру | 22 | 2 | - | 2 | 1 | 20 | Устный опрос, проверка практических заданий |
| 9. | Анализ текстов на естественном языке | 20 | 2 | - | 2 | 1 | 18 | Устный опрос, проверка практических заданий |
| 10. | Визуализация результатов анализа данных | 11 | 1 | - | 1 | - | 10 | Устный опрос, проверка практических заданий |
|  | В целом по дисциплине | 144 | 20 | 4 | 16 | 4 | 128 | Курсовая работа |
|  | Итого в % |  |  |  |  | 20% |  |  |

**5.3. Содержание семинаров, практических занятий**

|  |  |  |
| --- | --- | --- |
| **Наименование тем (разделов) дисциплины** | **Перечень вопросов для обсуждения на семинарских, практических занятиях, рекомендуемые источники из разделов 8,9 (указывается раздел и порядковый номер источника)** | **Формы проведения занятий** |
| Тема 1. Информационные технологии анализа данных | * Входной контроль. * Изучение технологического стека анализа данных, построенного на базе языка программирования Python.   *Рекомендуемые источники: 8.5; 8.8; 8.9; 8.10;* | Индивидуальное выполнение заданий, групповой разбор результатов выполнения заданий |
| Тема 2.  Введение в машинное обучение и знакомство с технологиями машинного обучения | * Решение задач машинного обучения с использованием библиотеки Scikit-Learn. * Оценка качества результатов: тестирование, разбиение набора данных для обучения и тестирования, кроссвалидация. * Расчет метрик качества результатов.   *Рекомендуемые источники: 8.5; 8.8; 8.9; 8.10;* | Индивидуальное выполнение заданий, групповой разбор результатов выполнения заданий |
| Тема 3.  Методы обучения с учителем | * Решение задач машинного обучения с использованием метрических методов классификации при помощи библиотеки Scikit-Learn. * Решение задач машинного обучения с использованием решающих списков и деревьев при помощи библиотеки Scikit-Learn. * Решение задач машинного обучения с использованием байесовских классификаторов при помощи библиотеки Scikit-Learn. * Решение задач машинного обучения с использованием метода опорных векторов при помощи библиотеки Scikit-Learn.   *Рекомендуемые источники: 8.5; 8.8; 8.9; 8.10;* | Индивидуальное выполнение заданий, групповой разбор результатов выполнения заданий |
| Тема 4.  Методы восстановления регрессии | * Решение задач восстановления регрессии при помощи библиотеки Scikit-Learn. * Решение задач машинного обучения с использованием логистической регрессии при помощи библиотеки Scikit-Learn.   *Рекомендуемые источники: 8.5; 8.8; 8.9; 8.10;* | Индивидуальное выполнение заданий, групповой разбор результатов выполнения заданий |
| Тема 5.  Обучение без учителя, кластеризация | * Решение задач кластеризации при помощи библиотеки Scikit-Learn (алгоритм k средних (k-means), другие методы кластеризации). * Решение задач снижения размерности при помощи библиотеки Scikit-Learn (метод главных компонент).   *Рекомендуемые источники: 8.5; 8.8; 8.9; 8.10;* | Индивидуальное выполнение заданий, групповой разбор результатов выполнения заданий |
| Тема 6. Коллаборативная фильтрация, отбор признаков | * Решение задач коллаборативной фильтрации на языке Python (корреляционные методы, латентные методы). * Реализация скользящего контроля на Python. Решение задачи отбора признаков на Python.   *Рекомендуемые источники: 8.5; 8.8; 8.9; 8.10;* | Индивидуальное выполнение заданий, групповой разбор результатов выполнения заданий |
| Тема 7. Генетические алгоритмы | * Выполнение аналитических задач с использованием технологий генетических алгоритмов.   *Рекомендуемые источники: 8.10;* | Выполнение практических заданий на компьютере, работа в группах |
| Тема 8.  Анализ сетевых структур | * Решение задач создания (в т.ч. загрузки из внешних источников и генерации) сетевых структур. Извлечение данных, выполнение запросов к графовой СУБД. Определение свойств сети, визуализация сети.   *Рекомендуемые источники: 8.1; 8.2;* | Выполнение практических заданий на компьютере, работа в группах |
| Тема 9.  Анализ текстов на естественном языке | * Токенезация, нормализация текста. Векторное представление текста, решение задач с помощью алгоритма TF-IDF.   *Рекомендуемые источники: 8.3; 8.4;* | Выполнение практических заданий на компьютере, работа в группах |
| Тема 10. Визуализация результатов анализа данных | * методы визуализации помогают улучшить понимание выявленных закономерностей; можно ли считать саму визуализацию инструментом анализа данных   *Рекомендуемые источники: 8.10;* | Выполнение практических заданий на компьютере, тематическая дискуссия |

**6.1. Перечень вопросов, отводимых на самостоятельное освоение дисциплины, формы внеаудиторной самостоятельной работы**

|  |  |  |
| --- | --- | --- |
| **Наименование тем (разделов) дисциплины** | **Перечень вопросов, отводимых на самостоятельное освоение** | **Формы внеаудиторной самостоятельной работы** |
| Тема1. Информационные технологии анализа данных | Знакомство с интерактивной оболочкой IPython notebook. Изучение принципов работы и применения для решения задач анализа данных и машинного обучения.  Знакомство с библиотеками numpy и pandas и решением базовых задач подготовительных операций для выполнения анализа данных с помощью этих библиотек. | Работа с литературой, работа с электронными источниками, разработка алгоритмов и программ |
| Тема 2.  Введение в машинное обучение и знакомство с технологиями машинного обучения | Расчет метрик качества Precision, Recal, F1, RUC. Критерии качества классификации: чувствительность и специфичность, ROC-кривая и AUC, точность и полнота, AUC-PR. | Работа с литературой, работа с электронными источниками, разработка алгоритмов и программ |
| Тема 3.  Методы обучения с учителем | Решение задач машинного обучения с использованием метрических методов классификации при помощи библиотеки Scikit-Learn (сравнение решения задач с различными параметрами и видами метрических алгоритмов).  Решение задач машинного обучения с использованием решающих списков и деревьев при помощи библиотеки Scikit-Learn (сравнение решения задач с различными параметрами и видами алгоритмов решающих списков и деревьев).  Решение задач машинного обучения с использованием байесовских классификаторов при помощи библиотеки Scikit-Learn (сравнение решения задач с различными параметрами и видами байесовских алгоритмов).  Решение задач машинного обучения с использованием метода опорных векторов при помощи библиотеки Scikit-Learn (сравнение решения задач с различными параметрами и видами алгоритмов метода опорных векторов). | Работа с литературой, работа с электронными источниками, разработка алгоритмов и программ |
| Тема 4.  Методы восстановления регрессии | Решение задач восстановления регрессии при помощи библиотеки Scikit-Learn с штрафами L1 и L2, Elastic Net.  Решение задач машинного обучения с использованием логистической регрессии при помощи библиотеки Scikit-Learn. | Работа с литературой, работа с электронными источниками, разработка алгоритмов и программ |
| Тема 5.  Обучение без учителя, кластеризация | Решение задач кластеризации при помощи библиотеки Scikit-Learn (использование методов, альтернативных алгоритму k средних).  Решение задач снижения размерности при помощи библиотеки Scikit-Learn (использование методов, альтернативных методу главных компонент). | Работа с литературой, работа с электронными источниками, разработка алгоритмов и программ |
| Тема 6. Коллаборативная фильтрация, отбор признаков | Решение задач коллаборативной фильтрации на языке Python корреляционными методами с различными метриками.  Тестирование альтернативных алгоритмов решения задачи отбора признаков. | Работа с литературой, работа с электронными источниками, разработка алгоритмов и программ |
| Тема 7.  Генетические алгоритмы | Генетические алгоритмы в прикладных задачах.  Преимущества и ограничения технологий генетических алгоритмов. | Работа с литературой, работа с электронными источниками |
| Тема 8.  Анализ сетевых структур | Идентификация сетей, имеющих структуру тесного мира  Визуализация сетей имеющих привязку узлов на плоскости | Работа с литературой, работа с электронными источниками |
| Тема 9.  Анализ текстов на естественном языке | Классификация тематического корпуса текстов | Работа с литературой, работа с электронными источниками |
| Тема 10. Визуализация результатов анализа данных | Современные методы визуализации аналитической информации  Краткое описание данных на основе визуализации | Работа с литературой, работа с электронными источниками, разработка алгоритмов и программ |

**6.2. Перечень вопросов, заданий, тем для подготовки к текущему контролю**

***Примерная тематика контрольных работ***

1. Визуализация результатов анализа данных
2. Анализ текстов на естественном языке
3. Коллаборативная фильтрация, отбор признаков
4. Методы обучения без учителя,
5. Кластеризация
6. Методы обучения с учителем

***Примерные вопросы к контрольным работам***

1. Какие виды активационных функций используются в искусственной нейронной сети типа RBF (сеть с радиальной базисной функцией)?
2. С помощью какого представления карты Кохонена можно понять сколько нейронов не стали победителями ни для одного образца?
3. Какие связи присутствуют в искусственной нейронной сети типа MLP (многослойный персептрон)?

***Примеры заданий контрольных работ***

**Задание 1:**

На основе предложенного массива данных о ценах на 1, 2 и 3-х комнатные квартиры в разных округах г. Москвы постройте необходимые визуальные представления, и ответьте на следующие вопросы

1. Как различаются между собой округа по следующим параметрам цен на квартиры: средняя, медиана, размах двух центральных квартилей, отличие средней цены первого и последнего дециля совокупности?
2. Как различается количественное предложение квартир с разным числом комнат по округам?

**Задание 2:**

На основе предложенного массива данных о продажах в супермаркете (покупательские корзины) выполните поиск шаблонов типа «ассоциация». На основе полученных данных выявите наиболее значимые с точки зрения возможности практического применения правила (5-7 правил). Обоснуйте ваш выбор на основе значений поддержки и достоверности.

***Примерная тематика курсовой работы***

1. Визуализация результатов анализа данных для набора данных биржевых торгов.
2. Визуализация результатов анализа данных для набора данных о корзинах покупок.
3. Визуализация результатов анализа данных для набора данных о продажах недвижимости.
4. Решение одной из задач анализа текстов на естественном языке для корпуса текстов на русском языке.
5. Решение одной из задач анализа текстов на естественном языке для корпуса текстов на английском языке.
6. Коллаборативная фильтрация для набора данных о покупках в интенет-магазине.
7. Коллаборативная фильтрация для набора данных об использовании хэш-тегов.
8. Решение задачи кластеризации для набора данных о покупках автомобилей.
9. Решение задачи кластеризации для набора данных о демографических показателях.
10. Решение задачи кластеризации для набора данных о продажах недвижимости.

**7. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине**

Перечень компетенций с указанием индикаторов их достижения в процессе освоения образовательной программы содержится в разделе 2 «Перечень планируемых результатов освоения образовательной программы с указанием индикаторов их достижения и планируемых результатов обучения по дисциплине».

**Типовые контрольные задания или иные материалы, необходимые для оценки индикаторов достижения компетенций, умений и знаний**

***Профиль «ИТ-сервисы и технологии обработки данных в экономике и финансах»***

**2019, 2020 год приема**

|  |  |  |
| --- | --- | --- |
| **Код компетенций** | **Наименование компетенций** | **Примеры заданий для оценки индикаторов достижения компетенций** |
| **ПКП-2** | Способность применять методы внедрения и эксплуатации корпоративных информационных систем в сфере экономики и финансов | **1.Демонстрирует знание назначения и функционал типовых модулей корпоративных информационных систем, основные этапы и методы их внедрения и эксплуатации**  **Задание 1.**  Опишите основные этапы внедрения CRM-системы и модули системы.  **2.Владеет методологией внедрения и эксплуатации приложений в сфере экономики и финансов на платформе корпоративных информационных систем.**  **Задание 2.**  Опишите основные функциональные модули типовой CRM-системы. |
| **ПКП-4** | Способность применять технологии моделирования и анализа процессов в сфере экономики и финансов | **1.Демонстрирует знания в области теории и методологии моделирования и анализа в сфере экономики и финансов.**  **Задание 1.**  Опишите принципиальную логику модели макроэкономической динамики основных фондов (модель Солоу).  **2.Строит математические модели в сфере экономики и финансов**  **Задание 2.**  Предложите варианты развития модели макроэкономической динамики основных фондов (модели Солоу). |

**2017, 2018 год приема**

|  |  |  |
| --- | --- | --- |
| **Код компетенций** | **Наименование компетенций** | **Примеры заданий для оценки сформированности компетенций** |
| ПК-7 | Способность проводить описание прикладных процессов и информационного обеспечения решения прикладных задач | Предложите технологический стек и архитектуру для решения задачи коллаборативной фильтрации с производительностью не менее 5000 t/s. |
| ПК-8 | Способность программировать приложения и создавать программные прототипы решения прикладных задач | Реализуйте решение задачи коллаборативной фильтрации для типичного интернет-магазина. |
| ПК-14 | Способность осуществлять ведение базы данных и поддержку информационного обеспечения решения прикладных задач | Опишите процесс решения задачи информационного обеспечения базы данных CRM-системы. |
| ПК-23 | Способность применять системный подход и математические методы в формализации решения прикладных задач | Предложите математическую модель макроэкономической динамики основных фондов. |
| ПКП-2 | Способность применять методы внедрения и эксплуатации корпоративных информационных систем в сфере экономики и финансов | Предложите технологический стек и архитектуру для решения задачи коллаборативной фильтрации с производительностью не менее 5000 t/s. |
| ПКП-4 | Способность применять технологии моделирования и анализа процессов в сфере экономики и финансов | Используя диаграммы языка UML опишите набор вариантов использования CRM-системы. |

***Профиль «Высокопроизводительные технологии в цифровой экономике»***

**2019, 2020 год приема**

|  |  |  |
| --- | --- | --- |
| **Код компетенций** | **Наименование компетенций** | **Примеры заданий для оценки индикаторов достижения компетенций** |
| ОПК-4 | Способен участвовать в разработке стандартов, норм и правил, а также технической документации, связанной с профессиональной деятельностью | **1.Разрабатывает техническую документацию для программных компонентов и информационных систем.**  **Задание 1.**  Предложите структуру технической документации для типовой CRM-системы.  **2. Работает со стандартами, в том числе адаптирует стандарты для специфических требований разрабатываемого программного обеспечения**.  **Задание 2.**  Перечислите стандарты, которые могут использоваться при описания технического задания на разработку ПО. |
| ПКП-7 | Способность провести статистический анализ данных, в том числе разнородных и данных большого объема | **1. Демонстрирует методы анализа статистических данных, рассчитывает показатели качества статистического анализа данных в зависимости от объема выборки.**  **Задание 1.**  Оценить выборочную дисперсию для рада данных, описывающих стоимость биржевого актива.  **2. Разрабатывает алгоритмы для обработки больших данных.**  **Задание 2.**  Опишите основной принцип построения алгоритма с использованием принципа map-reduce.  **3. Обосновывает выбор методов статистического анализа в зависимости от типов данных, в том числе нечисловых данных.**  **Задание 4.**  Перечислите известные Вам непараметрические методы статистического анализа. |

**2018 год приема**

|  |  |  |
| --- | --- | --- |
| **Код компетенций** | **Наименование компетенций** | **Примеры заданий для оценки сформированности компетенций** |
| ПК-14 | Способность осуществлять ведение базы данных и поддержку информационного обеспечения решения прикладных задач | Опишите процесс решения задачи информационного обеспечения базы данных CRM-системы. |
| ПК-20 | Способность осуществлять и обосновывать выбор проектных решений по видам обеспечения информационных систем | Перечислите известные Вам виды типовых проектных решений для CRM-систем. |
| ПКП-7 | Способность провести статистический анализ данных, в том числе разнородных и данных большого объема | Перечислите известные Вам непараметрические методы статистического анализа. |

***Примеры практико-ориентированных (ситуационных) заданий***

1. Проведите анализ новостного потока CNN (включая текстовые и видео сообщения) и провести статистический анализ влияния потока на биржевые показатели ключевых мировых рынков. Предложить подход к оценке качества и тестированию модели. Привести методы визуализации входных данных и результатов модели.
2. Предложите технологический стек и архитектуру для решения задачи коллаборативной фильтрации с производительностью не менее 5000 t/s. Предложить подход к оценке качества и тестированию модели. Привести методы визуализации входных данных и результатов модели.
3. Предложить пример матричного разложения для частично определённой матрицы 5 на 4, содержащей оценки по пятибалльной шкале. Показать, как с помощью разложения можно определить оценку для неопределенной ячейки матрицы. Дать содержательную интерпретацию задачи. Предложить подход к оценке качества и тестированию модели. Привести методы визуализации входных данных и результатов модели.
4. Решить задачу коллаборативной фильтрации на предложенных исходные данных (в т.ч. их матричном представлении). Корреляционные методы user-based, item-based. Различные меры сходства субъектов и объектов. Предложить подход к оценке качества и тестированию модели. Привести методы визуализации входных данных и результатов модели.
5. Для задачи классификации красных и синих объектов, предложенных на рисунке. Предложить графический вид трех моделей классификации: недообученной, переобученной и нормально обученной. Обосновать характеристики данные моделям. Предложить подход к оценке качества и тестированию модели. Привести методы визуализации входных данных и результатов модели.

**Примерные вопросы для подготовки к экзамену**

1. Понятие машинного обучения. Оценка качества результатов машинного обучения.
2. Виды шкал признаков. Подготовка и очистка данных для машинного обучения.
3. Понятия переобучения и недообучения. Диагностика переобучения и подходы к борьбе с ним.
4. Задача классификации и задача регрессии. Постановка, обзор подходов к решению.
5. Метод логистической регрессии.
6. Метод опорных векторов.
7. Метод линейной регрессии.
8. Метод градиентного спуска.
9. Метод случайного леса и метод решающих деревьев.
10. Задача кластеризации. Постановка и формализация.
11. Задача понижения размерности. Постановка и применение.
12. Перекрестная проверка (кросс-валидация). Описание метода и применение.
13. Задача обнаружения аномалий.
14. Закон Ципфа и стоп-слова в тексте
15. Языковые модели, n-граммы, вероятностная модель с допущением Маркова
16. Расстояние Левенштейна, алгоритм Вагнера-Фишера
17. Модель перцептрона. Проблема классификации для однослойной ИНС.
18. Механизм обратного распространения ошибки.
19. Метод стохастического градиентного спуска.
20. История развития методов построения ИНС. Принципы глубокого обучения.
21. Проблема переобучения ИНС и методы регуляризации ИНС.
22. Нормализация по мини-батчам.
23. Усовершенствованные методы градиентного спуска.
24. Kеras: общие принципы, технологический стек, последовательный и функциональный стиль описания моделей.
25. ТеnsorFlow: описание архитектуры вычислений.
26. Принцип работы сверточных ИНС.
27. Преобразование размерности данных при использовании сверточных сетей и операций пулинга.
28. Типичная архитектура сети распознавания изображений с помощью сверточных ИНС.
29. Организация глубокого обучения на небольших наборах изображений.
30. Специализированные базы данных фиксирующие лексические и семантические отношения между словами.
31. Дистрибутивная семантика, матрица совместной встречаемости и меры близости векторов.
32. Представление слов в виде векторов малой размерности с помощью сингулярного разложения.
33. Принцип работы Word2Vec, описание алгоритма CBOW
34. Принцип работы Word2Vec, описание алгоритма Skip-gram

***Примеры экзаменационных билетов***

**Билет№**

На основе предложенного массива данных о ценах на 1, 2 и 3-х комнатные квартиры в разных округах г. Москвы постройте необходимые визуальные представления, и ответьте на следующие вопросы:

1. Как различаются между собой округа по следующим параметрам цен на квартиры: средняя, медиана, размах двух центральных квартилей, отличие средней цены первого и последнего дециля совокупности?

2. Как различается количественное предложение квартир с разным числом комнат по округам?

**Билет №**

1. На основе предложенного массива данных о продажах в супермаркете (покупательские корзины) выполните поиск шаблонов типа «ассоциация».

2. На основе полученных данных выявите наиболее значимые с точки зрения возможности практического применения правила (5-7 правил). Обоснуйте ваш выбор на основе значений поддержки и достоверности.

**11. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень необходимого программного обеспечения и информационных справочных систем**

**11. 1. Комплект лицензионного программного обеспечения:**

1. Windows, Microsoft Office.

2. Антивирус ESET Endpoint Security

**11.2. Современные профессиональные базы данных и информационные справочные системы**

1. Информационно-правовая система «Гарант»

2. Информационно-правовая система «Консультант Плюс»

3. Электронная энциклопедия: <http://ru.wikipedia.org/wiki/Wiki>

4.Система комплексного раскрытия информации «СКРИН» -http://www.skrin.ru/

**11.3. Сертифицированные программные и аппаратные средства защиты информации**

**-** не предусмотрены