Федеральное государственное образовательное бюджетное учреждение

высшего образования

**«Финансовый университет при Правительстве**

**Российской Федерации»**

**(Финансовый университет)**

**Краснодарский филиал Финуниверситета**

Кафедра Математика и информатика

**ФОНД ОЦЕНОЧНЫХ СРЕДСТВ**

**по дисциплине**

**Построение и оценка моделей машинного обучения**

**Направление подготовки**

**01.04.02 – Прикладная математика и информатика**

**(магистратура)**

**Направленность программы магистратуры:**

**Анализ больших данных и машинное обучение в экономике и финансах**

**Форма обучения**

**Очная**

Составитель: **Калайдин Е.Н., профессор кафедры «Математика и информатика», д-р физ.-мат.наук**

*Рекомендовано Ученым советом Краснодарского филиала Финуниверситета*

*(протокол № 61 от 21.02.2023)*

*Одобрено кафедрой «Математика и информатика»*

*(протокол № 12 от 14.02.2023)*

**Краснодар**

**2023**

**Паспорт фонда оценочных средств**

Оценочные средства предназначены для контроля и оценки образовательных достижений обучающихся, освоивших программу учебной дисциплины «Построение и оценка моделей машинного обучения».

Фонд оценочных средств включает контрольные материалы для проведения текущего контроля и промежуточной аттестации.

1. **Описание показателей и критериев оценивания компетенций**

| Планируемые результаты освоения компетенции (индикатора достижения компетенции) | Уровень освоения | | | | Оценочное средство |
| --- | --- | --- | --- | --- | --- |
| «неудовлетворительно» | «удовлетворительно» | «хорошо» | «отлично» |
| **ПКН-4 Способность разрабатывать информационные системы и алгоритмы на основе математических методов и моделей, в том числе из области искусственного интеллекта, в решении профессиональных задач** | | | | | |
| Демонстрирует знания в области оригинальных алгоритмов и программной реализации | | | | | |
| **Знать:** теоретические основы в области оригинальных алгоритмов и программной реализации | Фрагментарное представление о теоретических основах в области оригинальных алгоритмов и программной реализации | Неполные представления о теоретических основах в области оригинальных алгоритмов и программной реализации | Сформированные, но содержащие отдельные пробелы представления о теоретических основах в области оригинальных алгоритмов и программной реализации | Сформированные систематические представления о теоретических основах в области оригинальных алгоритмов и программной реализации | Вопросы для оценки знаний и умений, практико-ориентированные задания, тесты |
| **Уметь:** демонстрировать знания в области оригинальных алгоритмов и программной реализации. | Фрагментарное умение демонстрировать знания в области оригинальных алгоритмов и программной реализации. | Несистематическое применение умений демонстрировать знания в области оригинальных алгоритмов и программной реализации. | В целом успешное, но содержащее отдельные пробелы умение демонстрировать знания в области оригинальных алгоритмов и программной реализации. | Сформированное умение демонстрировать знания в области оригинальных алгоритмов и программной реализации. | Вопросы для оценки знаний и умений, практико-ориентированные задания, тесты |
| Создает оригинальные алгоритмические и программные средства в решении профессиональных задач | | | | | |
| **Знать:** теоретические основы для создания оригинальных алгоритмических и программных средств в решении профессиональных задач. | Фрагментарное представление о теоретических основах для создания оригинальных алгоритмических и программных средств в решении профессиональных задач. | Неполные представления о теоретических основах для создания оригинальных алгоритмических и программных средств в решении профессиональных задач. | Сформированные, но содержащие отдельные пробелы представления о теоретических основах для создания оригинальных алгоритмических и программных средств в решении профессиональных задач. | Сформированные систематические представления о теоретических основах для создания оригинальных алгоритмических и программных средств в решении профессиональных задач. | Вопросы для оценки знаний и умений, практико-ориентированные задания, тесты |
| **Уметь:** создавать оригинальные алгоритмические и программные средства в решении профессиональных задач. | Фрагментарное умение создавать оригинальные алгоритмические и программные средства в решении профессиональных задач. | Несистематическое применение умений создания оригинальных алгоритмических и программных средств в решении профессиональных задач. | В целом успешное, но содержащее отдельные пробелы умения создавать оригинальные алгоритмические и программные средства в решении профессиональных задач. | Сформированное умение создавать оригинальные алгоритмические и программные средства в решении профессиональных задач. | Вопросы для оценки знаний и умений, практико-ориентированные задания, тесты |
| Демонстрирует знания в области современных интеллектуальных технологий | | | | | |
| **Знать:** теоретические основы в области современных интеллектуальных технологий. | Фрагментарное представление о теоретических основах в области современных интеллектуальных технологий. | Неполные представления о теоретических основах в области современных интеллектуальных технологий. | Сформированные, но содержащие отдельные пробелы представления о теоретических основах в области современных интеллектуальных технологий. | Сформированные систематические представления о теоретических основах в области современных интеллектуальных технологий. | Вопросы для оценки знаний и умений, практико-ориентированные задания, тесты |
| **Уметь:** демонстрировать знания в области современных интеллектуальных технологий. | Фрагментарное умение демонстрировать знания в области современных интеллектуальных технологий. | Несистематическое применение умений демонстрировать знания в области современных интеллектуальных технологий. | В целом успешное, но содержащее отдельные пробелы умение демонстрировать знания в области современных интеллектуальных технологий. | Сформированное умение демонстрировать знания в области современных интеллектуальных технологий. | Вопросы для оценки знаний и умений, практико-ориентированные задания, тесты |
| Выбирает интеллектуальные технологии и решает профессиональные задачи с их использованием | | | | | |
| **Знать:** различные интеллектуальные технологии. | Фрагментарное представление о различных интеллектуальных технологиях. | Неполные представления о различных интеллектуальных технологиях. | Сформированные, но содержащие отдельные пробелы представления о различных интеллектуальных технологиях. | Сформированные систематические представления о различных интеллектуальных технологиях. | Вопросы для оценки знаний и умений, практико-ориентированные задания, тесты |
| **Уметь:** выбирать интеллектуальные технологии и решать профессиональные задачи с их использованием. | Фрагментарное умение выбирать интеллектуальные технологии и решать профессиональные задачи с их использованием. | Несистематическое применение умений выбирать интеллектуальные технологии и решать профессиональные задачи с их использованием. | В целом успешное, но содержащее отдельные пробелы умение выбирать интеллектуальные технологии и решать профессиональные задачи с их использованием. | Сформированное умение выбирать интеллектуальные технологии и решать профессиональные задачи с их использованием. | Вопросы для оценки знаний и умений, практико-ориентированные задания, тесты |
| **ПКН-6 Способность анализировать и оценивать эффективность применения методов прикладной математики и информатики** | | | | | |
| Демонстрирует знание основных методов прикладной математики и информатики, применяемых в различных предметных областях. | | | | | |
| **Знать:** основные методы прикладной математики и информатики, применяемые в различных предметных областях. | Фрагментарное представление об основных методах прикладной математики и информатики, применяемых в различных предметных областях. | Неполные представления об основных методах прикладной математики и информатики, применяемых в различных предметных областях. | Сформированные, но содержащие отдельные пробелы представления об основных методах прикладной математики и информатики, применяемых в различных предметных областях. | Сформированные систематические представления об основных методах прикладной математики и информатики, применяемых в различных предметных областях. | Вопросы для оценки знаний и умений, практико-ориентированные задания, тесты |
| **Уметь:** демонстрировать знания основных методов прикладной математики и информатики, применяемые в различных предметных областях. | Фрагментарное умение различать основные методы прикладной математики и информатики, применяемые в различных предметных областях. | Несистематическое применение умений различать основные методы прикладной математики и информатики, применяемые в различных предметных областях. | В целом успешное, но содержащее отдельные пробелы умение различать основные методы прикладной математики и информатики, применяемые в различных предметных областях. | Сформированное умение различать основные методы прикладной математики и информатики, применяемые в различных предметных областях. | Вопросы для оценки знаний и умений, практико-ориентированные задания, тесты |
| Владеет методологией математического моделирования для решения профессиональных задач | | | | | |
| **Знать:** методологию математического моделирования для решения профессиональных задач. | Фрагментарное представление о методологии математического моделирования для решения профессиональных задач. | Неполные представления о методологии математического моделирования для решения профессиональных задач. | Сформированные, но содержащие отдельные пробелы представления о методологии математического моделирования для решения профессиональных задач. | Сформированные систематические представления о методологии математического моделирования для решения профессиональных задач. | Вопросы для оценки знаний и умений, практико-ориентированные задания, тесты |
| **Уметь:** владеть методологией математического моделирования для решения профессиональных задач. | Фрагментарное умение владеть методологией математического моделирования для решения профессиональных задач. | Несистематическое применение умений владеть методологией математического моделирования для решения профессиональных задач. | В целом успешное, но содержащее отдельные пробелы умение владеть методологией математического моделирования для решения профессиональных задач. | Сформированное умение владеть методологией математического моделирования для решения профессиональных задач. | Вопросы для оценки знаний и умений, практико-ориентированные задания, тесты |

**2****. Задания или иные материалы, необходимые для оценки знаний и умений, характеризующих формирование компетенций в процессе освоения ОП ВО**

***2.1 Вопросы для оценки знаний и умений, характеризующих формирование компетенций***

|  |  |  |
| --- | --- | --- |
| Шифр компетенции | Вопросы | Правильный ответ |
| ПКН-4 | 1. В каком году был создан первый искусственный интеллект? | 1950 |
| 2. Какие модели и методы используются для обработки естественного языка? | Нейронные сети, статистические модели, методы машинного обучения |
| 3. Какие модели и методы используются для компьютерного зрения? | CNN, GAN, RNN |
| 4. Какая модель обучения с подкреплением основана на методе Монте-Карло? | Модель Монте-Карло |
| 5. Как называется функция, которая используется для оценки качества модели в задачах классификации? | Функция качества классификации |
| 6. Что такое гиперпараметры? | Параметры, определяющие сложность модели |
| 7. С чем может быть связано переобучение модели? | Низкая обобщающая способность модели |
| 8. Какая проблема может возникнуть при работе с данными в машинном обучении? | Проблема несбалансированных классов |
| 9. Какой из методов предобработки наиболее эффективен для решения задачи классификации? | Стандартизация |
| 10. К какой категории методов оптимизации относится метод градиентного спуска? | Нелинейным |
| ПКН-6 | 11. Что позволяет теорема Байеса? | Учитывать априорные распределения вероятностей при обучении модели |
| 12. Какие задачи решаются с помощью глубокого обучения? | Классификация и регрессия |
| 13. Какую библиотеку можно использовать для анализа данных и выполнения статистических операций? | Pandas |
| 14. В чем состоит особенность модели дерева решений? | Построение иерархической структуры решений |
| 15. Какая модель используется для предсказания непрерывных целевых переменных? | Модель линейной регрессии |
| 16. Для чего используются ядерные методы в машинном обучении? | Аппроксимация нелинейных функций |
| 17. Для чего используется метод DBSCAN (Density-Based Spatial Clustering of Applications with Noise)? | Нахождение K-ближайших соседей и выделение кластеров |
| 18. Как выбираются локальные веса в локально-взвешенном сглаживании? | На основе расстояния до соседей |
| 19. Что такое бэггинг? | Метод ансамблевого обучения |
| 20. Для чего используется ARIMA (Autoregressive Integrated Moving Average)? | Прогнозирование временных рядов |

***2.2 Практико-ориентированные задания***

|  |  |  |
| --- | --- | --- |
| Шифр компетенции | Практико-ориентированные задания | Правильный ответ |
| ПКН-4 | 1. В базе данных имеется 10 000 записей. Сколько информации (в битах) необходимо для кодирования одной записи, если на каждый атрибут выделяется по 3 бита? | 29 |
| 2. В задаче регрессии имеется 500 точек данных. Сколько параметров требуется настроить в линейной регрессии, если каждая точка представлена двумя координатами (x, y)? | 1000 |
| 3. Имеется набор данных для классификации, состоящий из 50 000 примеров. Сколько раз необходимо выполнить итерацию обучения алгоритма k-ближайших соседей (k-NN), если в каждой итерации используется 10% набора данных для проверки? | 10 |
| 4. Ваша задача разработать систему рекомендаций для электронной коммерции. На основе какой фильтрации будет оптимально это сделать? | Коллаборативной |
| 5. Перед вами стоит задача построить модель прогнозированных временных рядов. Какие модельные формы вы будете для этого использовать? | ARIMA, LSTM |
| 6. Дан набор данных, состоящий из последовательностей действий пользователей на веб-сайте. Требуется классифицировать эти действия, то есть определить, какие из них являются целевыми, а какие – нет. Что для этого нужно сделать? | Определить цели, собрать данные, создать модель классификации и внедрить её на веб-сайте |
| 7. Доктор хочет построить модель, прогнозирующую возможность развития осложнений у пациента после операции - все ли будет благополучно в течение нескольких следующих месяцев или нет. К какому типу относится данная задача? | Классификация |
| 8. Астроном хочет построить модель, которая сможет разбить известные науке звезды на группы по их характеристикам, чтобы лучше изучить их особенности. К какому типу относится данная задача? | Кластеризация |
| 9. Заполните пропуск: «Для выпуклой функции .... указывает сторону наискорейшего убывания» | Градиент |
| ПКН-6 | 10. Что может являться критерием останова в градиентном спуске? | Достижение определенной точности |
| 11. Рассмотрим два объекта: у первого отступ линейного классификатора равен -10, у второго -1000. Как соотносятся два этих объекта? | Второй объект находится дальше от разделяющей поверхности, чем первый |
| 12. Рассмотрим пользователя социальной сети как объект в задаче машинного обучения. Что может являться задачей классификации? | Предсказание пола пользователя |
| 13. Ансамлевые методы обычно делают более качественные предсказания, чем отдельные алгоритмы (у ансамблей ниже ошибка на тестовой выборке). А какой показатель всегда ухудшается при ансамблировании? | Скорость выполнения предсказаний |
| 14. Алгоритм бэггинг подразумевает выбор случайных частей данных, обучение алгоритма (например, решающего дерева) на каждой части и составление ансамбля из обученных алгоритмов. Для чего нужно обучать алгоритмы на разных частях данных? | Для выполнения несовпадающих предсказаний |
| 15. Метеоролог хочет построить модель, прогнозирующую температуру воздуха на завтра. К какому типу относится данная задача? | Прогнозирование |
| 16. Верно ли суждение, что элементами обучающей выборки являются объекты, характеристики которых являются значениями признаков | Да |
| 17. Рассмотрим признак “Образовательная программа” при анализе данных по студентам университета. Этот признак может принимать три значения: “Экономика”, “Математика”, “Философия”. Воспользуемся one-hot кодированием и заменим этот признак на три бинарных, которые будут соответствовать категориям в том порядке, в котором они перечислены выше. Как будет закодирован признак со значением “Философия”? | (0,0,1) |
| 18. Чему будет равен корень из среднеквадратичной ошибки для набора из 3 наблюдений, где отклонение предсказания линейной регрессии от реальных значений равны: -1, 2, 2? | 2 |

***2.3 Тесты***

|  |  |  |
| --- | --- | --- |
| Шифр компетенции | Тестовые задания | Правильный ответ |
| ПКН-4 | 1. Определите какой из методов машинного обучения используется для прогнозирования временных рядов.   1. Метод опорных векторов 2. Деревья решений 3. Метод k-ближайших соседей 4. Регрессия | c) |
| 2. Переобучение в контексте машинного обучения — это ...   1. Процесс подбора оптимальных параметров модели машинного обучения 2. Ситуация, когда модель идеально подходит под обучающую выборку, но плохо работает на новых данных 3. Потеря обобщающей способности модели 4. Невозможность подобрать оптимальные параметры модели | b) |
| 3. Выберите метрику оценки качества модели машинного обучения для задачи классификации, которая отражает долю верно классифицированных объектов обучающей выборки:   1. Точность 2. Полнота 3. F1-мера 4. ROC-AUC | a) |
| 4. Выберите какие из перечисленных характеристик могут быть использованы для оценки качества модели машинного обучения.   1. Количество деревьев в ансамблевой модели 2. Время обучения модели 3. Точность модели на обучающей и тестовой выборке 4. Среднее время обработки одного объекта | c) |
| 5. Определите, что из перечисленного может являться причиной переобучения модели машинного обучения.   1. Недостаточное количество данных для обучения 2. Избыточное количество гиперпараметров модели 3. Недостаточное количество слоев в нейронной сети 4. Избыточная сложность модели | d) |
| 6. Метод регуляризации — это ...   1. Добавление случайного шума в обучающую выборку 2. Отбрасывание некоторых объектов из обучающей выборки 3. Ограничение сложности модели для предотвращения переобучения 4. Увеличение объема обучающей выборки | c) |
| ПКН-6 | 7. Определите метод, который используется для обнаружения переобучения модели машинного обучения.   1. Кросс-валидация 2. Регуляризация 3. Градиентный бустинг 4. Ансамблирование моделей | a) |
| 8. Определите инструмент, который используется для визуализации процесса обучения нейронной сети и анализа ее производительности.   1. TensorBoard 2. Jupyter Notebook 3. PyCharm 4. Visual Studio Code | a) |
| 9. Определите метод, который используется для определения оптимального числа кластеров при кластеризации данных.   1. Метод локтя (Elbow Method) 2. Алгоритм DBSCAN 3. Алгоритм k-ближайших соседей 4. Алгоритм опорных векторов (SVM) | a) |
| 10. Определите метод машинного обучения, который обычно используется для решения задач регрессии.   1. Метод опорных векторов 2. Деревья решений 3. Метод k-ближайших соседей 4. Линейная регрессия | d) |
| 11. Определите, что показывает ROC-кривая.   1. Чувствительность и специфичность модели 2. Вероятность ложного срабатывания и пропуска события 3. Точность и полноту модели 4. Вероятность ошибки модели | b) |
| 12. Выберите один или несколько вариантов ответа, в которых используются ансамблевые модели машинного обучения.   1. Если модель работает хорошо на обучающих данных, но имеет низкую обобщающую способность 2. Если задача имеет высокие размерности входных данных и малое количество обучающих образцов 3. Если требуется повысить точность модели 4. Если результаты модели неустойчивы и подвержены влиянию случайных флуктуаций | c) |

**3.** **Методические материалы, определяющие процедуры оценивания знаний и умений, характеризующих степень сформированности компетенций**

Промежуточная аттестация проводится в соответствии с [Положением о проведении текущего контроля успеваемости и промежуточной аттестации обучающихся по программам бакалавриата и магистратуры в Финансовом университете](http://tsu.ru/upload/medialibrary/c07/prikaz-870.pdf) (Приказ №0557/о от 23.03.2017 г.)

Форма промежуточной аттестации по дисциплине – *зачет.*

**Критерии оценивания знаний и умений, характеризующих степень сформированности компетенций:**

- оценкой ***«зачтено»*** оценивается полное освоение компетенций по данной дисциплине. Оценка выставляется при получении обучающимся 86 и более баллов. При этом он:

**знает**: теоретические основы в области оригинальных алгоритмов и программной реализации; теоретические основы для создания оригинальных алгоритмических и программных средств в решении профессиональных задач; теоретические основы в области современных интеллектуальных технологий; различные интеллектуальные технологии; основные методы прикладной математики и информатики, применяемые в различных предметных областях; методологию математического моделирования для решения профессиональных задач;

**умеет**: демонстрировать знания в области оригинальных алгоритмов и программной реализации; создавать оригинальные алгоритмические и программные средства в решении профессиональных задач; демонстрировать знания в области современных интеллектуальных технологий; выбирать интеллектуальные технологии и решать профессиональные задачи с их использованием; различать основные методы прикладной математики и информатики, применяемые в различных предметных областях; владеть методологией математического моделирования для решения профессиональных задач.

- оценка **«*не зачтено»*** выставляется в том случае, если компетенции не освоены, ответы содержат существенные ошибки и обучающимся получено менее 50 баллов. При этом он:

**не знает**: теоретические основы в области оригинальных алгоритмов и программной реализации; теоретические основы для создания оригинальных алгоритмических и программных средств в решении профессиональных задач; теоретические основы в области современных интеллектуальных технологий; различные интеллектуальные технологии; основные методы прикладной математики и информатики, применяемые в различных предметных областях; методологию математического моделирования для решения профессиональных задач.

**не умеет**: демонстрировать знания в области оригинальных алгоритмов и программной реализации; создавать оригинальные алгоритмические и программные средства в решении профессиональных задач; демонстрировать знания в области современных интеллектуальных технологий; выбирать интеллектуальные технологии и решать профессиональные задачи с их использованием; различать основные методы прикладной математики и информатики, применяемые в различных предметных областях; владеть методологией математического моделирования для решения профессиональных задач.