

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ  
УЧРЕЖДЕНИЕ ИНКЛЮЗИВНОГО ВЫСШЕГО ОБРАЗОВАНИЯ

**«МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ ГУМАНИТАРНО-  
ЭКОНОМИЧЕСКИЙ УНИВЕРСИТЕТ»**

---

Факультет прикладной математики и информатики  
Кафедра прикладной математики и информатики по областям

«Утверждаю»

Зав. кафедрой



Петрунина Е.В.

«27» августа 2018 г.

**ФОНД ОЦЕНОЧНЫХ СРЕДСТВ  
ПО ДИСЦИПЛИНЕ**

**Математические методы машинного обучения**

образовательная программа направления подготовки

09.03.03 Прикладная информатика

блок ФТД. В.02 «Дисциплины (модули)», вариативная часть, факультативы

Профиль подготовки

**Прикладная информатика в менеджменте**

Квалификация (степень) выпускника

Бакалавр

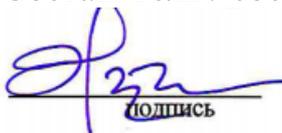
Форма обучения очная

Курс 4 семестр 7

Москва

2018

Составитель / составители: доцент кафедры ПМИИ по областям

  
Подпись

Никольский А.Е. «24»августа 2018 г.  
Ф.И.О. Дата

Фонд оценочных средств рассмотрен и одобрен на заседании кафедры ПМИИ по областям протокол № 1 от «27» августа 2018 г.

Зав. кафедрой  Петрунина Е.В. «27»августа 2018 г.  
Подпись Ф.И.О. Дата

Дополнения и изменения, внесенные в фонд оценочных средств, утверждены на заседании кафедры \_\_\_\_\_,  
протокол № \_\_\_\_ от « \_\_\_\_ » \_\_\_\_\_ 20\_\_ г.

Заведующий кафедрой \_\_\_\_\_ / Ф.И.О/

Дополнения и изменения, внесенные в фонд оценочных средств, утверждены на заседании кафедры \_\_\_\_\_,  
протокол № \_\_\_\_ от « \_\_\_\_ » \_\_\_\_\_ 20\_\_ г.

Заведующий кафедрой \_\_\_\_\_ / Ф.И.О/

Дополнения и изменения, внесенные в фонд оценочных средств, утверждены на заседании кафедры \_\_\_\_\_,  
протокол № \_\_\_\_ от « \_\_\_\_ » \_\_\_\_\_ 20\_\_ г.

Заведующий кафедрой \_\_\_\_\_ / Ф.И.О/

Дополнения и изменения, внесенные в фонд оценочных средств, утверждены на заседании кафедры \_\_\_\_\_,  
протокол № \_\_\_\_ от « \_\_\_\_ » \_\_\_\_\_ 20\_\_ г.

Заведующий кафедрой \_\_\_\_\_ / Ф.И.О/

## Содержание

1. Паспорт фонда оценочных средств.....
2. Перечень оценочных средств.....
3. Описание показателей и критериев оценивания результатов обучения на различных этапах формирования компетенций .....
4. Методические материалы, определяющие процедуры оценивания результатов обучения, характеризующих этапы формирования компетенций.....
5. Материалы для проведения текущего контроля и промежуточной аттестации.....

# 1. Паспорт фонда оценочных средств

по дисциплине «Математические методы машинного обучения»

Таблица 1.

№ п/п	Контролируемые разделы (темы), дисциплины <sup>1</sup>	Коды компетенций	Оценочные средства - наименование	
			текущий контроль	промежуточная аттестация
<i>7 семестр</i>				
1	Основные технологии, используемые в математических методах машинного обучения.	ОПК-3 ПК-7	Устный опрос, отчет по практической работе, реферат	<i>вопросы к зачету</i>
2.	Анализ многомерных данных. Корреляционные и причинно-следственные связи	ОПК-3 ПК-7	Устный опрос, отчет по практической работе	<i>вопросы к зачету / Зачет</i>

Таблица 2.

Перечень компетенций:

Код компетенции	Содержание компетенции
ОПК-3	Способность использовать основные законы естественнонаучных дисциплин и современные информационно-коммуникационные технологии в профессиональной деятельности
ПК-7	Способность проводить описание прикладных процессов и информационного обеспечения решения прикладных задач

<sup>1</sup> Наименование раздела (темы) берется из рабочей программы дисциплины.

## 2. Перечень оценочных средств<sup>2</sup>

Таблица 3.

№	Наименование оценочного средства	Характеристика оценочного средства	Представление оценочного средства в ФОС
1	Устный опрос	Средство контроля усвоения учебного материала темы, раздела или разделов дисциплины, организованное как учебное занятие в виде собеседования преподавателя с обучающимися.	Вопросы по темам/разделам дисциплины
2	Практические работы	Средство контроля усвоения учебного материала темы, раздела или разделов дисциплины, состоящее в индивидуальном выполнении обучающимся практических заданий для оценки полученных знаний, умений и владений компетенциями, формируемыми по данной дисциплине.	Задания для выполнения практических работ
3	Зачет		Вопросы к зачету

<sup>2</sup> Указываются оценочные средства, применяемые в ходе реализации рабочей программы данной дисциплины.

### 3. Описание показателей и критериев оценивания результатов обучения на различных этапах формирования компетенций

При проведении текущего контроля успеваемости студентов по учебной дисциплине ФТД.В.02 «Математические методы машинного обучения» используются следующие критерии оценок:

#### 3.1. Критерии оценки устного опроса

Устный опрос используется для текущего контроля успеваемости обучающихся по дисциплине в качестве проверки результатов освоения терминологии.

Каждому студенту выдается свой собственный, узко сформулированный вопрос.

Ответ должен быть четким и кратким, содержащим все основные характеристики описываемого понятия, института, категории.

#### Описание критериев и шкалы оценивания устного опроса

Критерий оценивания	Оценка
Выставляется обучающемуся, который подготовил ответ на предложенный вопрос, активно участвует в дискуссии, высказывает собственное мнение, представляет наглядный материал	Отлично
Выставляется обучающемуся, который подготовил ответ на предложенный вопрос, но неактивном участии в дискуссии	Хорошо
Выставляется обучающемуся, который частично подготовил ответ на предложенный вопрос, неактивно участвовал в дискуссии	Удовлетворительно
Выставляется обучающемуся в случае его неготовности к занятию	Неудовлетворительно

#### 3.2. Критерии оценки заданий в форме практических работ

Практические работы используются для текущего контроля успеваемости обучающихся по дисциплине в качестве проверки результатов освоения учебного материала.

Каждому студенту выдается своё собственное задание.

Отчет должен быть четким и кратким, содержащим все основные характеристики описываемого понятия, института, категории.

#### Описание критериев и шкалы оценивания заданий в форме практических работ

Критерий оценивания	Оценка
Выставляется обучающемуся, который подготовил отчет, в котором делаются самостоятельные выводы,	Отлично

дается аргументированная критика и самостоятельный анализ фактического материала на основе глубоких знаний литературы по данной теме	
Выставляется обучающемуся, проявившему полное и знание учебного материала, но нет должной степени самостоятельности	Хорошо
Выставляется обучающемуся, проявившему знания основного учебного материала в объеме, необходимом для последующего обучения и предстоящей практической деятельности, но в основном обладающему необходимыми знаниями и умениями для их устранения при корректировке со стороны преподавателя	Удовлетворительно
Выставляется обучающемуся, обнаружившему существенные пробелы в знании основного учебного материала, допустившему принципиальные ошибки при применении теоретических знаний, которые не позволяют ему продолжить обучение или приступить к практической деятельности без дополнительной подготовки по данной дисциплине	Неудовлетворительно

### 3.3. Критерии оценки зачета (зачета с оценкой)

В ходе ответа обучающийся должен показать сформированность компетенции (или компетенций) по дисциплине.

Результаты ответа определяются оценками «зачтено (отлично)», «зачтено (хорошо)», «зачтено (удовлетворительно)», «незачтено (неудовлетворительно)».

Зачет представляет собой форму промежуточного контроля знаний по дисциплине. Он проводится в устной форме. Каждому обучающемуся выдается два теоретических вопроса и одна задача.

На подготовку обучающегося отводится 30 минут.

#### Описание критериев и шкалы оценивания зачета (зачета с оценкой)

Показатели	Максимальная оценка в баллах
1-й вопрос	30
2-й вопрос	30
Задача	40

0-50 баллов	51-70	71-85	86-100
Незачтено (неудовлетворительно)	Зачтено (удовлетворительно)	Зачтено (хорошо)	Зачтено (отлично)

Для оценки уровня освоения дисциплин, профессиональных модулей (их составляющих) устанавливаются следующие соответствие:

«отлично» - высокий уровень освоения;

«хорошо», «удовлетворительно» - достаточный уровень освоения;

«неудовлетворительно» - низкий уровень освоения.

**Таблица 4.**

Код компетенции	Уровень освоения компетенции	Показатели достижения компетенции	Критерии оценивания результатов обучения
		Знает	
ОПК-3	Недостаточный уровень. Оценка «незачтено»	ОПК-3. З-1.	Не знает значительной части материала курса, не способен самостоятельно выделять главные положения в изученном материале дисциплины
	Базовый уровень. Оценка «зачтено»		Знает не менее 50 % основного материала курса, однако испытывает затруднения в его применении
	Средний уровень. Оценка «зачтено»		Знает основную часть материала курса, способен применить изученный материал на практике, испытывает незначительные затруднения в решении задач
	Высокий уровень. Оценка «зачтено»		Показывает глубокое знание и понимание материала, способен применить изученный материал на практике
		Умеет	
	Базовый уровень	ОПК-3. У-1.	Умеет воспроизвести не менее 50 % основного материала курса, однако испытывает затруднения при решении практических задач
	Средний уровень		Умеет решать стандартные профессиональные задачи с применением полученных знаний, испытывает незначительные затруднения в решении задач
	Высокий уровень		Умеет решать стандартные профессиональные задачи с применением полученных знаний, показывает глубокое знание и понимание материала, способен решить задачу при изменении формулировки
		Владеет	
	Базовый уровень	ОПК-3. В-1.	Владеет навыками теоретического и экспериментального исследования объектов профессиональной деятельности, усвоил основное содержание материала дисциплины, но имеет пробелы в усвоении материала. Имеет несистематизированные знания основных разделов дисциплины.
	Средний уровень		Владеет навыками теоретического и экспериментального исследования объектов профессиональной деятельности, способен самостоятельно выделять главные положения в изученном материале. Испытывает незначительные затруднения в решении задач.
	Высокий уровень		Свободно владеет навыками теоретического и экспериментального исследования, показывает глубокое знание и понимание изученного материала

ПК-7		Знает	
	Недостаточный уровень. Оценка «незачтено»	ПК-7. З-1.	Не знает значительной части материала курса, не способен самостоятельно выделять главные положения в изученном материале дисциплины
	Базовый уровень. Оценка «зачтено»		Знает не менее 50 % основного материала курса, однако испытывает затруднения в его применении
	Средний уровень. Оценка «зачтено»		Знает основную часть материала курса, способен применить изученный материал на практике, испытывает незначительные затруднения в решении задач
	Высокий уровень. Оценка «зачтено»		Показывает глубокое знание и понимание материала, способен применить изученный материал на практике
		Умеет	
	Базовый уровень	ПК-7. У-1.	Умеет воспроизвести не менее 50 % основного материала курса, однако испытывает затруднения при решении практических задач
	Средний уровень		Умеет решать стандартные профессиональные задачи с применением полученных знаний, испытывает незначительные затруднения в решении задач
	Высокий уровень		Умеет решать стандартные профессиональные задачи с применением полученных знаний, показывает глубокое знание и понимание материала, способен решить задачу при изменении формулировки
		Владеет	
	Базовый уровень	ПК-7. В-1.	Владеет навыками теоретического и экспериментального исследования объектов профессиональной деятельности, усвоил основное содержание материала дисциплины, но имеет пробелы в усвоении материала. Имеет несистематизированные знания основных разделов дисциплины.
	Средний уровень		Владеет навыками теоретического и экспериментального исследования объектов профессиональной деятельности, способен самостоятельно выделять главные положения в изученном материале. Испытывает незначительные затруднения в решении задач.
Высокий уровень	Свободно владеет навыками теоретического и экспериментального исследования, показывает глубокое знание и понимание изученного материала		

#### **4. Методические материалы, определяющие процедуры оценивания результатов обучения**

По видам заданий приводится описание того, каким образом необходимо выполнить данное задание, способы и механизмы его выполнения, выбор номера варианта и др. Примеры методических материалов, определяющих процедуру оценивания результатов обучения, характеризующих этапы формирования компетенций:

- Кейсовые технологии как средство формирования компетенций
- Методические указания по разработке оценочных средств
- Разработка и применение деловых игр
- Иные методические материалы, определяющие процедуры оценивания результатов обучения в ходе реализации рабочей программы дисциплины

##### **Задания в форме устного опроса:**

Устный опрос используется для текущего контроля успеваемости обучающихся по дисциплине в качестве проверки результатов освоения материала. Каждому студенту выдается свой собственный, узко сформулированный вопрос. Ответ должен быть четким и кратким, содержащим все основные характеристики описываемого понятия. В своем ответе студент должен показать умения прослеживать причинно-следственные связи и навыки рассуждений и доказательства.

##### **Задания в форме практических работ:**

Практическая работа представляет собой контрольное мероприятие по учебному материалу каждой темы (раздела) дисциплины, состоящее в индивидуальном выполнении обучающимся практических заданий для оценки полученных знаний, умений и владений компетенциями, формируемыми по данной дисциплине.

Выполнение практических работ является средством текущего контроля успеваемости обучающихся по дисциплине и может включать в себя следующие типы заданий: задания типового вида и задания творческого характера. По результатам выполнения практических заданий обучающие оформляют отчеты, содержащие анализ полученных результатов и выводы.

#### **5. Материалы для проведения текущего контроля и промежуточной аттестации**

##### **Задания в форме устного опроса**

##### **Семестр 7**

##### **Раздел 1. Основные технологии, используемые в методах машинного обучения.**

- 1) Основные понятия и определения в методах машинного обучения.
- 2) История развития методов машинного обучения.
- 3) Методологическая база методов машинного обучения.
- 4) Типы и способы представления методов машинного обучения
- 5) Базовые алгоритмы решения задач машинного обучения
- 6) Основные программно-информационные ресурсы методов машинного обучения

##### **Раздел 2. Анализ многомерных данных. Корреляционные и причинно-следственные связи**

- 1) Анализ многомерных данных.

- 2) Корреляционные и причинно-следственные связи. Корреляция признаков и структура данных.
- 3) Регрессия. Метод наименьших квадратов.
- 4) Теорема Гаусса-Маркова. Обобщенный метод наименьших квадратов.
- 5) Многомерная регрессия. Особенности построения регрессии по многомерным данным.
- 6) Множественная линейная регрессия, ее преимущества и недостатки.
- 7) Кластеризация. Кластеризация как классификация без учителя.
- 8) Меры сходства и меры различия образов.
- 9) Метод K средних.
- 10) Метод ISODATA.
- 11) Метод FOREL.

### **Контролируемые компетенции: ОПК-3, ПК-7**

**Оценка компетенций осуществляется в соответствии с Таблицей 4.**

### **Задания в форме практических работ**

#### **Семестр 7**

Практическая работа 1: «Автоматизация конструирования признаков»

Цель работы: приобрести практический опыт автоматизации конструирования признаков.

Задание: найти оптимальное представление набора данных с помощью метода главных компонент (МГК) и критериев отбора компонент.

Шаги:

1 Сгенерировать набор данных произвольной размерности  $d$ , например, методом «`sklearn.datasets.make_blobs`».

2 Получить дисперсии значений компонент. Это можно сделать с помощью класса «`sklearn.decomposition.PCA`» и атрибута «`explained_variance_`».

3 Определить пороговые значения дисперсии компонент с помощью:

- правила Кайзера;
- правила сломанной трости;
- «метода локтя».

4 Построить диаграммы, например, с помощью метода «`matplotlib.pyplot.bar`», отметить на диаграмме пороговые значения по каждому критерию.

5 Применить МГК с количеством компонент, полученным на основе одного из критериев.

Отчет должен содержать следующие файлы:

1 Сгенерированный набор данных в формате `.csv`.

2 Диаграмма, полученная на шаге 4

3 Преобразованный набор данных, полученный на шаге 5, в формате `.csv`.

Практическая работа 2: «Автоматизация выбора модели и оптимизации гиперпараметров»

Цель работы: приобрести практический опыт автоматизации классификации.

Задание: разработать автоматизированный классификатор.

Шаги:

1 Сгенерировать набор данных для классификации, например, методом «sklearn.datasets.make\_blobs».

2 Выбрать не менее 3 алгоритмов классификации, например, из библиотеки «sklearn».

3 Выбрать целевую функцию. Если была использована библиотека «sklearn», то функции можно выбрать из модулю «sklearn.metrics».

4 Для каждого выбранного классификатора реализовать автоматическую оптимизацию гиперпараметров методами поиска по решетке и случайного поиска с кросс-валидацией. При использовании «sklearn» можно воспользоваться методами «model\_selection.GridSearchCV» и «model\_selection.RandomizedSearchCV».

5 Реализовать автоматизированный выбор классификатора из сформированного набора с оптимизированными гиперпараметрами на основе выбранной целевой функции. 6 Применить выбранный классификатор.

Отчет должен содержать следующие файлы:

1 Сгенерированный набор данных в формате .csv.

2 Текстовый файл, в котором указаны:

- выбранные классификаторы;

- выбранная целевая функция;

- список полученных в результате оптимизации значений гиперпараметров для каждого классификатора и значения целевой функции.

3 Результат классификации в формате .csv.

**Контролируемые компетенции: ОПК-3, ПК-7**

**Оценка компетенций осуществляется в соответствии с Таблицей 4.**

**Тематика рефератов**

**Семестр 7**

1. Работа с типами данных в языке Python.

2. Введение в массивы библиотеки NumPy.

3. Выполнение вычислений над массивами библиотеки NumPy.

4. Операции над данными в библиотеке Pandas.

5. Визуализация с помощью библиотеки Matplotlib.

6. Библиотека Scikit-Learn.

7. Смеси Гауссовых распределений.

8. Ядерная оценка плотности распределения.

9. Метод опорных векторов. Оптимальная разделяющая гиперплоскость.

10. Случаи линейной делимости и отсутствия линейной делимости.

Кусочно-линейная функция потерь.

11. Задача квадратичного программирования и двойственная задача.

12. Понятие опорных векторов. Линейные методы классификации.

13. Градиентные методы. Линейный классификатор, связь с методом максимума правдоподобия.

14. Метод стохастического градиента и частные случаи: адаптивный линейный элемент ADALINE, перцептрон Розенблатта, правило Хэбба.

15. Метрические методы классификации. Метод ближайших соседей и его обобщения.

16. Постановка задач обучения по прецедентам.

17. Типы задач: классификация, регрессия, прогнозирование, кластеризация.

Примеры прикладных задач.

18. Основные понятия: модель алгоритмов, метод обучения, функция потерь и функционал качества.
19. Методика экспериментального исследования и сравнения алгоритмов на модельных и реальных данных.
20. Полигон алгоритмов классификации.
21. CRISP-DM — межотраслевой стандарт ведения проектов интеллектуального анализа данных.
22. Понятие логической закономерности. Решающие списки и деревья. Объединение в решающие леса.
23. Сингулярное разложение, метод главных компонент.
24. Наивная байесовская классификация.
25. Машинное обучение с учителем и обучение без учителя.

### **Контролируемые компетенции: ОПК-3, ПК-7**

**Оценка компетенций осуществляется в соответствии с Таблицей 4.**

### **Вопросы к зачету**

#### **Семестр 7**

1. Основные понятия. Определение предмета машинного обучения. Примеры задач и областей приложения. Образы и признаки.
2. Типы задач предсказания. Регрессия. Таксономия. Классификация. Типы ошибок классификации. Обобщающая способность классификатора.
3. Принцип минимизации эмпирического риска. Недообучение. Переобучение. Статистический, нейросетевой и структурно-лигвистический подходы к распознаванию образов.
4. Структура типичной системы распознавания образов. Цикл построения системы распознавания образов.
5. Классификация. Общие принципы. Этапы классификации. Алгоритмы обучения классификаторов с учителем и без учителя. Дискриминантный анализ. Геометрическая интерпретация задачи классификации.
6. Проективный подход. Метрики в пространстве признаков. Евклидово расстояние. Расстояние Махаланобиса. Ошибки первого и второго рода. Чувствительность и избирательность.
7. Кривая мощности критерия классификации. ROC-кривые. Проверка классификатора. Проверка тестовой выборкой. Перекрестная проверка. Оценка информативности признаков.
8. Основные методы машинного обучения. Байесовская классификация. Условная вероятность. Формула полной вероятности. Формула Байеса. Статистическое распознавание образов. Задача классификации спама. Критерий отношения правдоподобия. Байесовский риск. Критерий Байеса.
9. Критерий максимального правдоподобия. Многоклассовые байесовские классификаторы. Байесовские классификаторы для нормально распределенных классов при различной структуре матрицы ковариации
10. Оценивание функций распределения. Параметрическое оценивание. Метод максимума правдоподобия. Байесовское оценивание. Непараметрическое оценивание. Распознавание рукописных цифр с помощью наивного байесовского
11. Деревья решений. Основные понятия. Классы решаемых задач: описание данных, классификация, регрессия. Общий алгоритм построения дерева решений. Критерии выбора наилучшего атрибута: прирост информации, относительный прирост информации, индекс Гини.
12. Правила остановки разбиения дерева. Обрезание дерева. Алгоритм ID3. Переобучение деревьев решений. Обработка непрерывных атрибутов. Обучение на

данных с пропусками. Программное обеспечение для построения деревьев решений. Распознавание спамовых писем с помощью деревьев решений

13. Анализ многомерных данных. Корреляционные и причинно-следственные связи. Корреляция признаков и структура данных.

14. Метод главных компонент как декомпозиция матрицы данных. Матрица счетов. Матрица нагрузок. Матрица ошибок. Объясненная и остаточная вариация в данных. Графическая интерпретация метода главных компонент. Критерии выбора количества главных компонент. Понижение размерности признакового пространства методом главных компонент

15. Регрессия. Метод наименьших квадратов. Теорема Гаусса-Маркова. Обобщенный метод наименьших квадратов. Рекурсивный метод наименьших квадратов. Анализ регрессионных остатков.

16. Многомерная регрессия. Особенности построения регрессии по многомерным данным. Множественная линейная регрессия, ее преимущества и недостатки.

17. Кластеризация. Кластеризация как классификация без учителя. Меры сходства и меры различия образов. Метод K средних. Метод ISODATA. Метод FOREL.

18. Графовые методы. Иерархическая кластеризация. Агломеративные и разделяющие алгоритмы кластеризации. Дендрограммы.

19. Нейронные сети. Предпосылки возникновения нейросетей. Перцептрон Розенблатта. Многослойный перцептрон. Карты Кохонена. Сети Хопфилда. Методы обучения нейросетей. Метод опорных векторов.

20. Машинное обучение и теория Вапника-Червоненкиса. Принцип структурной минимизации риска. Метод опорных векторов. Политика назначения штрафов.

**Контролируемые компетенции: ОПК-3, ПК-7**

**Оценка компетенций осуществляется в соответствии с Таблицей 4.**