

**Министерство науки и высшего образования  
Российской Федерации  
Министерство науки, высшего образования и  
инноваций Кыргызской Республики**

**Государственное образовательное  
учреждение  
высшего профессионального образования  
Кыргызско-Российский Славянский  
университет имени первого президента  
Российской Федерации Б.Н. Ельцина**

**Естественно-технический факультет**

**Кафедра Информационных и вычислительных  
технологий**

**Фонд  
оценочных средств**  
по дисциплине «Машинное обучение и анализ  
данных»

Уровень высшего образования

**МАГИСТРАТУРА**

Направление подготовки

09.04.04 - РФ, 710400 - КР Программная инженерия  
(код и наименование направления подготовки)

Разработка программно-информационных систем  
(профиля) образовательной программы)

Квалификация  
*магистр*

Фонд оценочных средств предназначен для контроля знаний обучающихся по направлению подготовки 09.04.04 – РФ, 710400 – КР «Программная инженерия» по дисциплине «Машинное обучение и анализ данных».

Фонд оценочных средств рассмотрен и утвержден на заседании кафедры Информационных и вычислительных технологий

Заведующий кафедрой  
д.т.н., проф.



Лыченко Н.М.

*Исполнители (разработчики):*

*К.т.н., доцент каф. ИВТ Верзунов С.Н.*

*К.т.н., доцент каф. ИВТ Манжикова С.Ц.*



---

СОГЛАСОВАНО:

И.о. декана ЕТФ Комарцов Н.М.



**Раздел 1. Перечень компетенций, с указанием этапов их формирования в процессе освоения дисциплины**

Формируемые компетенции	Планируемые результаты обучения по дисциплине, характеризующие этапы формирования компетенций	Виды оценочных средств / шифр раздела в данном документе
<p><b>ПК-4: Владение навыками создания программного обеспечения для анализа, распознавания и обработки информации, систем цифровой обработки сигналов</b></p>	<p><b>Знать:</b></p> <ul style="list-style-type: none"> <li>• основные понятия и методы машинного обучения и анализа данных;</li> <li>• постановки задач классификации, регрессии, кластеризации, тематического моделирования;</li> <li>• основы вероятностного подхода, принципы регуляризации, переобучения и валидации;</li> <li>• метрики качества (Accuracy, Precision, Recall, F1, ROC-AUC и др.);</li> <li>• современные программные средства разработки и анализа данных на Python.</li> </ul>	<p><b>Блок А</b> – задания репродуктивного уровня:</p> <ul style="list-style-type: none"> <li>– Тест;</li> <li>– Теоретический опрос.</li> </ul>
	<p><b>Уметь:</b></p> <ul style="list-style-type: none"> <li>• формулировать прикладные задачи машинного обучения;</li> <li>• выбирать и применять адекватные методы и модели;</li> <li>• проводить предобработку данных и настройку гиперпараметров;</li> <li>• оценивать качество моделей и интерпретировать результаты.</li> </ul>	<p><b>Блок В</b> – задания реконструктивного уровня:</p> <ul style="list-style-type: none"> <li>– Практические задания;</li> <li>– Отчет по практическим работам.</li> </ul>
	<p><b>Владеть:</b></p> <ul style="list-style-type: none"> <li>• практическими навыками разработки программных решений для анализа и распознавания данных;</li> <li>• современными технологиями построения ML-пайплайна;</li> <li>• навыками экспериментальной валидации моделей и подготовки аналитических отчетов.</li> </ul>	<p><b>Блок С</b> – задания практико-ориентированного и/или исследовательского уровня:</p> <ul style="list-style-type: none"> <li>– Комплексные практические задания;</li> <li>– Мини-исследование (по заданию преподавателя).</li> </ul>

## Раздел 2. Технологическая карта дисциплины

### Машинное обучение и анализ данных

Курс 1, семестр 2, Количество ЗЕ – 4, Отчетность – зачет с оценкой

Название модулей дисциплины согласно РПД	Контроль	Форма контроля	зачетный минимум	зачетный максимум	График контроля
1. Введение в машинное обучение и анализ данных	текущий	Сдача практической работы №1	5	8	4
	рубежный	Тест №1	5	8	4
2. Методы классификации: логистическая регрессия, SVM, метрики качества	текущий	Сдача практической работы №2	5	9	8
	рубежный	Тест №2	5	9	8
3. Ансамбли и многометочная классификация	текущий	Сдача практической работы №3	5	9	12
	рубежный	Тест №3	5	9	12
4. Современные модели классификации изображений. Кластеризация и анализ результатов	текущий	Сдача практической работы №4	5	9	16
	рубежный	Тест №4	5	9	16
<b>ВСЕГО за семестр</b>			<b>40</b>	<b>70</b>	
<b>Промежуточный контроль (зачет с оценкой)</b>		Письменный ответ + практическое задание	<b>20</b>	<b>30</b>	
<b>Семестровый рейтинг по дисциплине</b>			<b>60</b>	<b>100</b>	

## Раздел 3. Типовые контрольные задания и иные материалы, необходимые для оценки планируемых результатов обучения по дисциплине (оценочные средства)

### Блок А

#### Задания репродуктивного уровня (ЗНАТЬ)

##### 3.1 Примерный перечень тестовых заданий

1. Укажите корректное определение обучающей выборки и признакового пространства.
2. В чем связь принципа максимума правдоподобия и минимизации эмпирического риска?
3. Что характеризует переобучение модели?
4. Какие виды признаков применяются в ML (числовые, категориальные, бинарные и др.)?
5. Что такое регуляризация и зачем она используется?

6. Укажите верные утверждения о логистической регрессии.
7. Что такое ROC-кривая и AUC?
8. Какие предпосылки у наивного байесовского классификатора?
9. Что такое метод опорных векторов в линейно разделимом случае?
10. Чем отличаются bagging и boosting?
11. Что такое LSA, PLSA, LDA в тематическом моделировании?
12. В чем состоит идея EM-алгоритма?
13. Выберите корректные утверждения о методе k-средних.
14. Для чего применяют DBSCAN и как интерпретируются `eps`, `min_samples`?
15. Как соотносятся precision, recall, F1-score в задачах классификации?
16. Что такое data leakage и почему оно приводит к завышенной оценке качества модели?
17. В чем различие hold-out и k-fold cross-validation?
18. Почему для несбалансированных данных accuracy может быть недостаточной метрикой?
19. Чем отличаются macro-, micro- и weighted-усреднения метрик?
20. Что такое baseline-модель и зачем она нужна в эксперименте?
- 21. 3.2 Вопросы к теоретическому опросу (пример)**
22. Модель алгоритмов, функционал качества, метод обучения.
23. Вероятностная постановка задачи обучения.
24. Логистическая регрессия и калибровка вероятностей.
25. Метрики качества классификации, ROC-AUC.
26. SVM: линейный и нелинейный случаи.
27. Композиции алгоритмов: случайный лес, градиентный бустинг.
28. Тематическое моделирование: TF-IDF, LSA, LDA.
29. Кластеризация: k-means, EM, DBSCAN.
30. Переобучение и методы его предотвращения.
31. Практические аспекты подбора гиперпараметров и интерпретации результатов.

## **Блок В и С**

### **Практические задания (УМЕТЬ и ВЛАДЕТЬ)**

#### **Практическая работа 1. Обучение классификатора цифр**

##### **Задания:**

1. Загрузите набор данных (Digits или эквивалент).
2. Выполните предобработку данных.
3. Обучите базовый классификатор (kNN / Logistic Regression / SVM).
4. Оцените качество модели (accuracy, confusion matrix, classification report).
5. Выполните сравнение параметров и сделайте выводы.

#### **Практическая работа 2. Классификация изображений**

##### **Задания:**

1. Подготовьте выборку изображений.
2. Реализуйте извлечение признаков/базовый pipeline.
3. Обучите не менее двух моделей и выполните сравнительный анализ.

4. Постройте ROC-кривые и сравните AUC (бинарный сценарий или one-vs-rest).
5. Проанализируйте типичные ошибки классификации.

### **Практическая работа 3. Классификация по нескольким меткам (multi-label)**

#### **Задания:**

1. Подготовьте многометочную выборку.
2. Реализуйте one-vs-rest либо иной подход.
3. Вычислите micro/macro-метрики качества.
4. Проанализируйте ошибки модели.
5. Предложите и проверьте один способ улучшения качества.

### **Практическая работа 4. Современная модель классификации изображений**

#### **Задания:**

1. Используйте современный подход (в т.ч. с использованием Colab при наличии инфраструктуры).
2. Выполните обучение/дообучение модели.
3. Сравните результаты с базовым решением.
4. Выполните анализ переобучения и обобщающей способности модели.
5. Подготовьте отчет с анализом качества и ограничений.

### **Комплексные задания исследовательского уровня (Блок С)**

1. Постройте и сравните ансамбли (Random Forest vs Gradient Boosting) на одном датасете.
2. Выполните тематическое моделирование корпуса текстов (TF-IDF + LDA/LSA), интерпретируйте темы.
3. Исследуйте устойчивость модели к выбросам и дисбалансу классов, предложите меры компенсации.
4. Реализуйте кластеризацию (k-means и DBSCAN), проведите сравнение по внутренним метрикам.
5. Подготовьте структурированный аналитический отчет с обоснованием выбора метода.

### **Блок D (промежуточный контроль)**

#### **Вопросы для проверки уровня обученности ЗНАТЬ**

1. Основные понятия и характеристики машинного обучения и анализа данных.
2. Формализация задачи обучения и математическая постановка.
3. Переобучение и обобщающая способность.
4. Вероятностные методы классификации.
5. Логистическая регрессия и регуляризация.
6. ROC, AUC и другие метрики качества.
7. SVM: постановка, dual-форма, kernel trick.
8. Деревья решений и их редукция.
9. Ансамбли: bagging, boosting, random forest.
10. Кластеризация: k-means, EM, DBSCAN.
11. Тематическое моделирование: LSA/PLSA/LDA.
12. Практические аспекты выбора модели и валидации.
13. Типичные ошибки в ML-проектах (утечка данных, неверная метрика, смещение выборки).
14. Подходы к интерпретации результатов модели.

### **Задания для проверки уровня обученности УМЕТЬ и ВЛАДЕТЬ**

1. Постройте классификатор и сравните модели по выбранным метрикам.
2. Постройте ROC-кривую, вычислите AUC и интерпретируйте результат.
3. Выполните настройку гиперпараметров и сравните качество до/после.
4. Постройте кластеризацию и проанализируйте структуру данных.
5. Подготовьте краткий технический отчет по результатам экспериментов.
6. Выявите ошибки классификации и предложите обоснованные улучшения.

### **Пример**

#### **ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ ПРОМЕЖУТОЧНОГО КОНТРОЛЯ № \_\_\_\_**

1. **Вопросы для проверки уровня обученности ЗНАТЬ**
2. 1.1 Вероятностная постановка задачи обучения.  
1.2 Композиции алгоритмов: bagging и boosting.
3. **Задание для проверки уровня обученности УМЕТЬ и ВЛАДЕТЬ**
4. На предоставленном наборе данных обучить модель классификации, оценить качество (не менее 3 метрик), выполнить анализ ошибок и предложить одно улучшение.

#### **Раздел 4. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций**

##### **Описание показателей и критериев оценивания компетенций, описание шкал оценивания**

##### **Применяемые оценочные средства:**

- Сдача практических работ на практических занятиях в соответствии с технологической картой дисциплины (текущая и рубежная аттестация);
  - Тест;
  - Письменный опрос по экзаменационным билетам (промежуточная аттестация – зачет с оценкой).
- Все виды оценочных средств оцениваются в соответствии со шкалами оценивания.
- Устный опрос на практических занятиях по отдельным темам проводится в течение всего периода обучения дисциплине.
- Результаты опроса учитываются при оценивании практических работ.

##### **ШКАЛА ОЦЕНИВАНИЯ ТЕСТА (рубежный контроль)**

- В одном тестовом задании некоторое количество закрытых вопросов.
- К заданиям даются готовые ответы на выбор: один правильный и остальные неправильные.
- Обучающемуся необходимо помнить: в каждом задании с выбором одного правильного ответа правильный ответ должен быть.
- За каждый правильный ответ – 5 баллов.
- Определяется сумма набранных баллов.
- Вычисляется отношение набранных баллов к максимально возможному за тест. Отметка выражается в %.

### **ШКАЛА ОЦЕНИВАНИЯ ПРАКТИЧЕСКИХ (ЛАБОРАТОРНЫХ) РАБОТ (текущий контроль)**

- **85–100 %** – Демонстрирует полное понимание проблемы. Все требования, предъявляемые к заданию, выполнены.
- **70–84 %** – Демонстрирует значительное понимание проблемы. Все требования, предъявляемые к заданию, выполнены.
- **60–69 %** – Демонстрирует частичное понимание проблемы. Большинство требований, предъявляемых к заданию, выполнены.
- **31–60 %** – Демонстрирует небольшое понимание проблемы. Многие требования, предъявляемые к заданию, не выполнены.
- **0–30 %** – Демонстрирует непонимание проблемы и даже не было попытки решить задачу.

### **ШКАЛА ОЦЕНИВАНИЯ ПИСЬМЕННОГО ОПРОСА (промежуточный контроль – «ЗНАТЬ»)**

- Отметкой (**7–10 баллов**) оценивается ответ, который показывает прочные знания теоретических основ дисциплины, понимание и правильное применение терминологии, правильные ответы на 75–100% вопросов.
- Отметкой (**5–7 баллов**) оценивается ответ, который показывает знание теоретических основ дисциплины, но неполное понимание и не всегда правильное применение терминологии; даны правильные ответы на 50–74% вопросов, в ответах допущено некоторое количество неточностей.
- Отметкой (**3–4 балла**) оценивается ответ, свидетельствующий о знакомстве с некоторыми теоретическими основами дисциплины; даны правильные ответы на 25–49% вопросов, допущены неточности и ошибки.
- Отметкой (**2 балла**) оценивается ответ, обнаруживающий незнание теоретических основ дисциплины. Отмечается отсутствие логичности и последовательности в ответе. Менее 25% правильных ответов. Допущены серьезные ошибки в содержании ответа.
- Отметкой (**0–1 балл**) оценивается ответ, при котором студент демонстрирует непонимание поставленных вопросов, или нет ответа.

### **ШКАЛА ОЦЕНИВАНИЯ ПРАКТИЧЕСКИХ ЗАДАНИЙ (промежуточный контроль – «УМЕТЬ и ВЛАДЕТЬ»)**

- Отметкой (**8–10 баллов**) оценивается ответ, при котором студент правильно отвечает на поставленные вопросы, демонстрирует полное понимание проблемы. Все требования, предъявляемые к заданию, выполнены.
- Отметкой (**5–7 баллов**) оценивается ответ, при котором студент в основном правильно отвечает на поставленные вопросы. Демонстрирует значительное понимание проблемы. Большинство требований, предъявляемых к заданию, выполнены.
- Отметкой (**2–4 балла**) оценивается ответ, при котором студент в основном неправильно отвечает на поставленные вопросы. Демонстрирует частичное или небольшое понимание проблемы. Многие требования, предъявляемые к заданию, не выполнены.
- Отметкой (**0–1 балл**) оценивается ответ, при котором студент демонстрирует непонимание проблемы или нет ответа и даже не было попытки решить задачу.

В экзаменационный билет включены два теоретических вопроса и практическое задание, соответствующие содержанию формируемых компетенций. Зачет с оценкой проводится в письменной форме. На ответ и решение задачи студенту отводится 80 минут. За ответ на теоретические вопросы студент может получить максимально 15 баллов, за выполнение практического задания — 15 баллов.

По итогам прохождения дисциплины и с учетом шкал оценивания все набранные в результате текущей, рубежной и промежуточной аттестаций баллы суммируются и выставляется оценка.

**Перевод баллов в оценку:**

- 85–100 баллов – «отлично»
- 70–84 баллов – «хорошо»
- 60–69 баллов – «удовлетворительно»
- менее 60 баллов – «неудовлетворительно»

**Раздел 5. Методические указания для обучающегося по освоению дисциплины и выполнению контрольных заданий**

**Рекомендации по организации самостоятельной работы студента**

**1. Советы по планированию и организации времени, необходимого для изучения дисциплины**

Рекомендуется следующим образом организовать время, необходимое для изучения дисциплины:

- Изучение конспекта лекции в тот же день, после лекции – 10–15 минут;
- Изучение конспекта лекции за день перед следующей лекцией – 10–15 минут;
- Изучение теоретического материала по учебному пособию и конспекту – 1 час в неделю;
- Подготовка к практическому занятию – 2–3 часа.

**Всего в неделю – 4 часа.**

**2. Описание последовательности действий студента**

Для понимания материала и качественного его усвоения рекомендуется следующая последовательность действий:

- После прослушивания лекции и окончания учебных занятий, при подготовке к занятиям следующего дня, сначала просмотреть и обдумать текст лекции, прослушанной сегодня (10–15 минут).
- При подготовке к лекции следующего дня просмотреть текст предыдущей лекции, подумать о возможной теме следующей лекции (10–15 минут).
- В течение недели выбрать время (2–3 часа) для работы с рекомендуемыми электронными учебными пособиями.
- При подготовке к практическим занятиям необходимо сначала прочитать основные понятия и подходы по теме домашнего задания. При выполнении упражнения или задачи нужно сначала понять, что требуется в задаче, какой теоретический материал использовать, наметить план решения.

**3. Рекомендации по использованию материалов учебно-методического комплекса**

Все рекомендуемые учебные пособия и методические материалы размещаются в ЭИОС/сетевых папках кафедры и на образовательной платформе дисциплины.