**Федеральное государственное образовательное бюджетное учреждение**

**высшего образования**

**«ФинансоВЫЙ УНИВЕРСИТЕТ**

**при Правительстве Российской Федерации»**

**(Финансовый университет)**

**Департамент анализа данных,  
принятия решений и финансовых технологий**

**Золотарева Екатерина Леоновна**

**РЕКОМЕНДАТЕЛЬНЫЕ СИСТЕМЫ**

Рабочая программа дисциплины

для студентов, обучающихся по направлению подготовки

01.04.02 «Прикладная математика и информатика»,

направленность программы магистратуры «Анализ больших данных и машинное обучение в экономике и финансах»

Москва 2018 год

**Федеральное государственное образовательное бюджетное учреждение**

**высшего образования**

**«ФинансоВЫЙ УНИВЕРСИТЕТ**

**при Правительстве Российской Федерации»**

**(Финансовый университет)**

**Департамент анализа данных,  
принятия решений и финансовых технологий**

|  |  |
| --- | --- |
| СОГЛАСОВАНО  \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  (наименование организации)  \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  (должность представителя работодателя)  \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ ФИО  (подпись представителя работодателя)  «\_\_\_\_» \_\_\_\_\_\_\_\_\_\_\_\_\_ 2018 г. | УТВЕРЖДАЮ  Ректор Финуниверситета \_\_\_\_\_\_\_\_\_\_\_\_М.А. Эскиндаров  24.04.2018 г. |

**Золотарева Екатерина Леоновна**

**РЕКОМЕНДАТЕЛЬНЫЕ СИСТЕМЫ**

Рабочая программа дисциплины

для студентов, обучающихся по направлению подготовки

01.04.02 «Прикладная математика и информатика»,

направленность программы магистратуры «Анализ больших данных и машинное обучение в экономике и финансах»

*Рекомендовано Ученым советом факультета*

*«Прикладная математика и информационные технологии»*

(*протокол № 04 от 17.04.2018 г.*)

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

*(протокол № 11от 17.04.2018 г.)*

Москва 2018 год

**УДК 519 (073)**

**ББК 22.16я73**

**С 60**

**Рецензент:** Феклин В.Г., к.ф.-м.н., доцент Департамента анализа данных, принятия решений и финансовых технологий

**С 60   Золотарева Е.Л.**

**Рекомендательные системы.** Рабочая программа дисциплины для студентов, обучающихся по направлению подготовки 01.04.02 «Прикладная математика и информатика» (направленность программы магистратуры «Анализ больших данных и машинное обучение в экономике и финансах»). – М.: Финуниверситет, Департамент анализа данных, принятия решений и финансовых технологий, 2018. – 22 с.

В рабочей программе дисциплины определены ее цель, место в структуре ОП, требования к результатам освоения дисциплины, содержание программы, тематика практических занятий, формы самостоятельной работы, оценочные средства для текущего контроля и промежуточной аттестации, учебно-методическое и информационное обеспечение.

***Учебное издание***

***Золотарева Екатерина Леоновна***

**РЕКОМЕНДАТЕЛЬНЫЕ СИСТЕМЫ**

*Рабочая программа дисциплины*

Компьютерный набор и верстка Е.Л. Золотарева

Формат 60×90/16. Гарнитура Times New Roman  
Усл.п.л. 1,5. Изд. № - 2018. Тираж экз.

Заказ № \_\_\_\_\_\_\_\_\_\_

Отпечатано в Финансовом университете

© Золотарева Е.Л., 2018

© Финансовый университет,2018

СОДЕРЖАНИЕ

[1. Наименование дисциплины 4](#_Toc515633621)

[2. Перечень планируемых результатов освоения образовательной программы с указанием индикаторов их достижения, соотнесенных с планируемыми результатами обучения по дисциплине 4](#_Toc515633622)

[3. Место дисциплины в структуре образовательной программы 8](#_Toc515633623)

[4. Объем дисциплины(модуля) в зачетных единицах и в академических часах с выделением объема аудиторной (лекции, семинары) и самостоятельной работы обучающихся 8](#_Toc515633624)

[5. Содержание дисциплины, структурированное по темам (разделам) дисциплины с указанием их объемов (в академических часах) и видов учебных занятий 9](#_Toc515633625)

[5.1. Содержание дисциплины 9](#_Toc515633626)

[5.2. Учебно – тематический план 11](#_Toc515633627)

[5.3. Содержание семинаров, практических занятий 13](#_Toc515633628)

[6. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине 14](#_Toc515633629)

[6.1. Перечень вопросов, отводимых на самостоятельное освоение дисциплины, формы внеаудиторной самостоятельной работы 14](#_Toc515633630)

[6.2. Перечень вопросов, заданий, тем для подготовки к текущему контролю (согласно таблице 2) 16](#_Toc515633631)

[7. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине 18](#_Toc515633632)

[7.1 перечень компетенций с указанием индикаторов их достижения в процессе освоения образовательной программы 19](#_Toc515633633)

[7.2 типовые контрольные задания или иные материалы, необходимые для оценки индикаторов достижения компетенций, умений и знаний; 19](#_Toc515633634)

[7.3 Приказы, распоряжения ректората о контроле уровня освоения дисциплин и сформированности компетенций студентов. 20](#_Toc515633635)

[8. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины 20](#_Toc515633636)

[9. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины 21](#_Toc515633637)

[10. Методические указания для обучающихся по освоению дисциплины 22](#_Toc515633638)

[11. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень необходимого программного обеспечения и информационных справочных систем (при необходимости). 22](#_Toc515633639)

[12. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине. 22](#_Toc515633640)

# 1. Наименование дисциплины

Дисциплина «Рекомендательные системы».

Основная цель данной дисциплины – обучить студентов инструментарию построения рекомендательных систем, их оценкой и применением в социальной и финансовой сфере.

# 2. Перечень планируемых результатов освоения образовательной программы с указанием индикаторов их достижения, соотнесенных с планируемыми результатами обучения по дисциплине

Дисциплина «Рекомендательные системы» обеспечивает инструментарий формирования следующих компетенций:

| Код компетенции | Наименование компетенции | Индикаторы достижения компетенции | Результаты обучения (владения, умения и знания), соотнесенные с компетенциями/индикаторами достижения компетенции |
| --- | --- | --- | --- |
| ДКМП-3 | Умение строить и оценивать модели машинного обучения в прикладных задачах | - | **Знать:**   * Общие принципы построения моделей машинного обучения * Подходы к валидации и кросс-валидации моделей * Основные метрики оценки качества моделей, их преимущества и недостатки   **Уметь:**   * Построить модель классификации/ регрессии * Рассчитать соответствующие метрики качества * Интерпретировать результат и сделать выводы о направлении доработки модели   **Владеть:**   * Методологией построения и оценки моделей машинного обучения * Практическими навыками разработки моделей машинного обучения |
| ДКПМ-5 | Умение строить рекомендательные системы и применять их к решению практических задач | - | **Знать:**   * Основные понятия в области рекомендательных систем * Принципы построения неперсонализированных и слабоперсонализированных систем * Принципы построения контентно-ориентированных систем * Принципы построения систем коллаборативной фильтрации * Принципы построения продвинутых и гибридных систем * Специфические метрики оценки качества рекомендательных систем, их преимущества и недостатки   **Уметь:**   * построить рекомендательную систему на основе статистики продаж и демографической (или иной неперсонализированной информации) * выделить атрибуты товаров/ пользователей на основе анализа нечисловой информации и построить контентно-ориентированную систему * построить систему коллаборативной фильтрации на основе алгоритмов «Пользователь-Пользователь» или «Товар-Товар» * аппроксимировать матрицу рейтингов матрицей меньшего ранга и использовать ее для гибридной рекомендательной системы * отобрать, рассчитать и интерпретировать релевантные метрики качества в зависимости от бизнес-целей заказчика системы   **Владеть:**   * методологией построения и оценки рекомендательных систем * практическими навыками разработки рекомендательных |
| ДКМП-6 | Способность создавать ИТ-сервисы, основанные на анализе данных и машинном обучении |  | **Знать:**   * Специализированное ПО и основные инструменты для решения задач машинного обучения   **Уметь:**   * Создавать, настраивать и оценивать модели в специализированном ПО * Настраивать обмен данными между специализированным ПО, источником данных и конечным потребителем   **Владеть:**   * Методологией создания ИТ-сервисов, основанных на анализе данных в машинном обучении * Практическими навыками создания ИТ-сервисов, основанных на анализе данных в машинном обучении |

# 

# 3. Место дисциплины в структуре образовательной программы

Дисциплина «Рекомендательные системы» входит в модуль обязательных дисциплин магистерской программы «Анализ больших данных и машинное обучение в экономике и финансах» направления подготовки 01.04.02 «Прикладная математика и информатика» и относится к классу дисциплин, ориентированных на формирование профессиональных компетенций выпускников в области разработки прикладных систем машинного обучения, в частности, рекомендательных систем.

В процессе изучения дисциплины происходит овладение основными компьютерными технологиями обработки больших данных и машинного обучения, которые будут положены в основу формирования профессиональных компетенций выпускника направления «Прикладная математика и информатика». При этом студенты приобретают опыт практического использования изучаемых технологий в практических задачах, связанных с финансовыми технологиями.

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

Общая трудоёмкость дисциплины обучения составляет 3 зачетных единицы. Вид текущего контроля - контрольная работа в 6-м модуле, вид промежуточной аттестации – экзамен.

Таблица 1

|  |  |  |
| --- | --- | --- |
| **Вид учебной работы по дисциплине** | **Всего**  **(в з/е и часах)** | **Модуль 6**  **(в часах)** |
| **Общая трудоемкость дисциплины** | 3 з. е. (108 часов) | 108 |
| ***Контактная работа - Аудиторные занятия*** | 30 | 30 |
| *Лекции* | 10 | 10 |
| *Семинары, практические занятия* | 20 | 20 |
| ***Самостоятельная работа*** | 78 | 78 |
| Вид текущего контроля | экзамен | экзамен |
| Вид промежуточной аттестации | Контрольная работа | Контрольная работа |

# 5. Содержание дисциплины, структурированное по темам (разделам) дисциплины с указанием их объемов (в академических часах) и видов учебных занятий

## 5.1. Содержание дисциплины

**Тема 1. Введение.**

Понятие и история развития рекомендательных систем. Основная терминология: пользователи, товары (items), рейтинги, предпочтения и рекомендаций, атрибуты пользователей и товаров. Критерии классификации рекомендательных систем: сфера применения, цель, источник рекомендации, внешние факторы, уровень персонализации, надежность и степень конфиденциальности, интерфейс выдачи рекомендации, рекомендательный алгоритм. Примеры успешной реализации бизнес-моделей, основанных на использовании рекомендательных систем (Amazon.com, MovieLens).

**Тема 2.** **Неперсонализированные рекомендательные системы** Неперсонализированные и стереотипные рекомендательные системы. Использование описательной статистики при построении рекомендаций, предпочтения большинства. Демографические данные (возраст, пол, национальность, социальный статус, местонахождение) как средство персонализации рекомендаций. Исследование связей между товарами (кросс-продажи). Недостатки неперсонализированных и слабо-персонализированных систем.

**Тема 3.** **Контентно-ориентированные рекомендательные системы.**

Построение контентно-ориентированной рекомендательной системы. Фильтрация на основе контента. Персонализаций рекомендаций на основе предпочтения пользователя. Способы получения и обработки информации о предпочтениях пользователя: прямое анкетирование, анализ активности, выделение связи между предпочтениями пользователя и атрибутами товаров. Анализ ключевых слов для формирования атрибутов товаров. TDIDF анализ и его производные для выделения ключевых слов. Математическая модель рекомендательной системы. Построение и актуализация профилей пользователей, взвешивание/нормализация рейтингов. Представление атрибутов пользователей и товаров в векторной форме; расчет рекомендаций на основе оценки расстояния между векторами. Регрессионные/классификационные модели для предсказания рейтинга пользователей. Преимущества и недостатки контентно-ориентированного подхода.

**Тема 4. Коллаборативная фильтрация. Метод ближайших «соседей».**

Понятие коллаборативной фильтрации, использование рейтингов пользователей для построения рекомендаций без учета атрибутов. Сравнительный анализ с контентно-ориентированным подходом.

Метода ближайших «соседей» и иные методы для нахождения схожих пользователей и товаров. Меры оценки сходства пользователей, преимущества и недостатки.

Алгоритм «Пользователь-Пользователь» для построения рекомендации на основе оценок пользователей со схожими рейтингами. Математическая модель, предпосылки и ограничения. Использование нормализации, снижение «зашумленности» рейтингов.

Алгоритм «Товар-Товар»: рекомендации схожих товаров на основе связей между их рейтингами. Ограничения алгоритма «Пользователь-Пользователь: «пробелы» в данных, вычислительные емкость, необходимость актуализации рекомендаций. Двухшаговая реализация алгоритма «Товар-Товар»: определение сходства товаров на основе их рейтингов, построение предсказания рейтинга на основе рейтинга «соседей». Математическая модель, предпосылки и ограничения.

**Тема 5. Метрики оценки качества рекомендательных систем**

Типы метрик в зависимости от объекта оценки и бизнес-целей. Цели оценки моделей (выбор алгоритма, настройка, сравнение) и ограничения, определение «плохой» и «хорошей» рекомендации.

Базовые метрики оценки качества предсказанных рейтингов и рекомендаций. Метрики оценки точности/погрешности: MAE, RMSE, MSE. Метрики поддержки принятия решений: ROC AUC, полноста и точность. Поиск нерелевантных рекомендаций («reversals»).

Продвинутые метрики. Метрики, ориентированные на пользователя: охват (доля товаров, для которых имеется рекомендации /персонализированные рекомендации), удержание пользователей, степень принятия рекомендаций пользователем, степень персонализации рекомендации, продвижение непопулярных товаров, степень диверсификации рекомендаций и т.п. Метрики, основанные на рангах: MAP, nDCG, nDPM

Общие принципы обучения и валидации моделей: разделение тестовой и обучающей выборки, кросс-валидация, подбор гиперпараметров, проблема переобучения. Работа с пропущенными данными (холодный старт), устранение эффекта популярного товара и пр.

Онлайн оценка качества, оценка с помощью пользователей. Источники обратной связи. Эксперименты и A/B тесты. Выбор показателей оценки эффективности (количество просмотров, объем продаж, конверсия и т.п.). Этапы реализации A/B теста. Статистический анализ результата.

**Тема 6. Продвинутые и гибридные модели рекомендательных систем. Приближение матрицей меньшего ранга.**

Задача снижения размерности. Сингулярное разложение матрицы рейтингов (SVD) и приближение матрицей меньшего ранга. Использование градиентного спуска для приближенного разложения матрицы (FunkSVD). Математическая модель FunkSVD. Стохастические алгоритмы для разложения матриц.

Гибридные модели. Совмещение контентно-ориентированного подхода и коллаборативной фильтрации. Примеры реализаций.

## 5.2. Учебно – тематический план

Таблица 2

| №  п/п | **Наименование тем (разделов) дисциплины** | **Трудоемкость в часах** | | | | | | **Формы текущего контроля успеваемости** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Всего** | **Аудиторная работа** | | | | **Самостоятельная работа** |
| Общая, в т.ч.: | Лекции | Семинары, практические занятия | Занятия в интерактивных формах |
| 1. | Тема 1. Введение. | 3 | 1 | 1 | - | - | 2 | Контрольные вопросы по теории |
| 2. | Тема 2. Неперсонализированные рекомендательные системы | 9 | 3 | 1 | 2 | 2 | 6 | Лабораторные работы и контрольная работа.  (построение компьютерной системы обработки данных, анализ конкретной ситуации и обсуждение результатов) |
| 3. | Тема 3. Контентно- ориентированные рекомендательные системы | 14 | 4 | 2 | 2 | 2 | 10 |
| 4. | Тема 4. Коллаборативная фильтрация. Метод ближайших соседей. | 28 | 8 | 2 | 6 | 6 | 20 |
| 5. | Тема 5. Метрики оценки качества рекомендательных систем | 22 | 6 | 2 | 4 | 4 | 16 |
| 6. | Тема 6. Продвинутые и гибридные модели рекомендательных систем. Приближение матрицей меньшего ранга. | 32 | 8 | 2 | 6 | 6 | 24 |
|  | В целом по дисциплине | 108 | 30 | 10 | 20 | 20 | 78 | Согласно учебному плану: контрольная работа |
|  | Итого в % |  |  |  |  | 67% |  |  |

## 5.3. Содержание семинаров, практических занятий

Практические занятия направлены на интенсивное овладение практическими навыками построения рекомендательных систем, что отражено в форме их проведения: лабораторные работы и контрольная работа.

Таблица 3

| **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 2.***  Неперсонализированные рекомендательные системы | 1. Построение рекомендаций на основе описательной статистики с учетом демографических групп (на примере статистики ресторанов или фильмов) 2. Построение рекомендательной системы на основе кросс-продаж товаров (на примере статистики фильмов) с использованием элементов программирования   Рекомендуемые источники для самостоятельной работы: а) основная: 1; б) дополнительная: 3, 4; в) Интернет-ресурсы: 8-9 | Лабораторная работа № 1.  Лабораторная работа №2. |
| ***Тема 3.***  Контентно- ориентированные рекомендательные системы | 1. Построение рекомендаций на основе рейтингов, проставленных пользователями для товаров с известными атрибутами 2. Построение рекомендаций на основе рейтингов, проставленных пользователями для товаров с неизвестными атрибутами. Извлечение атрибутов на основе программной реализации TFIDF анализа   Рекомендуемые источники для самостоятельной работы: а) основная: 1,2; б) дополнительная: 1, 2; в) Интернет-ресурсы: 5-9 | Лабораторная работа № 3.  Лабораторная работа № 4. |
| ***Тема 4.***  Коллаборативная фильтрация. Метод ближайших соседей. | 1. Поиск ближайших «соседей»-пользователей/товаров на основе матрицы рейтингов (на примере фильмов) 2. Построение рекомендаций с помощью алгоритма «Пользователь-Пользователь» или «Товар-Товар» с использованием элементов программирования   Рекомендуемые источники для самостоятельной работы: а) основная: 1,2; б) дополнительная: 1, 2; в) Интернет-ресурсы: 5-9 | Лабораторная работа № 5.  Лабораторная работа № 6. |
| ***Тема 5.***  Метрики оценки качества рекомендательных систем | 1. Расчет числовых метрик качества построенных рекомендательных систем и интерпретация результатов   Рекомендуемые источники для самостоятельной работы: а) основная: 1,2; б) дополнительная: 1, 2, 3 в) Интернет-ресурсы: 4-9 | Лабораторная работа № 7.  Контрольная работа |
| ***Тема 6.***  Продвинутые и гибридные модели рекомендательных систем. Приближение матрицей меньшего ранга. | 1. Аппроксимация матрицы рейтингов матрицей меньшего ранга 2. Разработка и оценка рекомендательной системы с использованием SVD 3. Построение гибридной рекомендательной системы   Рекомендуемые источники для самостоятельной работы: а) основная: 1,2; б) дополнительная: 1, 2; в) Интернет-ресурсы: 6-9 | Лабораторная работа № 8.  Лабораторная работа № 10. |

# 6. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине

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

При изучении дисциплины « Рекомендательные системы» обязательными являются следующие формы самостоятельной работы:

Таблица 4

| **Наименование тем (разделов) дисциплины** | **Перечень вопросов, отводимых на самостоятельное освоение** | **Формы внеаудиторной самостоятельной работы** |
| --- | --- | --- |
| ***Тема 1.*** Введение | Примеры успешной реализации бизнес-моделей, основанных на использовании рекомендательных систем (Amazon.com, MovieLens). | * работа с текстом лекции, разбор вопросов по теме занятия; * изучение рекомендованных к занятию литературных источников; |
| ***Тема 2.***  Неперсонализированные рекомендательные системы | Стереотипные рекомендательные системы.  Недостатки неперсонализированных и слабо-персонализированных систем. | * работа с текстом лекции, разбор вопросов по теме занятия; * изучение рекомендованных к занятию литературных источников; * подготовка к лабораторным работам, настройка программного обеспечения; * разработка/отладка рекомендательной системы в соответствии с конкретными задачами лабораторной работы |
| ***Тема 3.***  Контентно- ориентированные рекомендательные системы | Способы получения и обработки информации о предпочтениях пользователя.  Расширения и альтернативы TFIDF-анализа | * работа с текстом лекции, разбор вопросов по теме занятия; * изучение рекомендованных к занятию литературных источников; * подготовка к лабораторным работам, настройка программного обеспечения; * разработка/отладка рекомендательной системы в соответствии с конкретными задачами лабораторной работы |
| ***Тема 4.***  Коллаборативная фильтрация. Метод ближайших соседей. | Использование нормализации, снижение «зашумленности» рейтингов.  Математическое обоснование (доказательство) используемых моделей | * работа с текстом лекции, разбор вопросов по теме занятия; * изучение рекомендованных к занятию литературных источников; * подготовка к лабораторным работам, настройка программного обеспечения; * разработка/отладка рекомендательной системы в соответствии с конкретными задачами лабораторной работы |
| ***Тема 5.***  Метрики оценки качества рекомендательных систем | Базовые метрики оценки качества предсказанных рейтингов и рекомендаций.  Общие принципы обучения и валидации моделей. | * работа с текстом лекции, разбор вопросов по теме занятия; * изучение рекомендованных к занятию литературных источников; * подготовка к лабораторным работам, настройка программного обеспечения; * разработка/отладка рекомендательной системы в соответствии с конкретными задачами лабораторной работы * повторение материалы и подготовка к контрольной работе |
| ***Тема 6.***  Продвинутые и гибридные модели рекомендательных систем. Приближение матрицей меньшего ранга. | Математическая модель FunkSVD. Стохастические алгоритмы для разложения матриц. | * работа с текстом лекции, разбор вопросов по теме занятия; * изучение рекомендованных к занятию литературных источников; * подготовка к лабораторным работам, настройка программного обеспечения; * разработка/отладка рекомендательной системы в соответствии с конкретными задачами лабораторной работы * повторение материала и подготовка к экзамену |

## 6.2. Перечень вопросов, заданий, тем для подготовки к текущему контролю (согласно таблице 2)

Перечень вопросов, задание и тем для подготовки к текущему контролю, а также тематика лабораторных работ определяется разделом 5 (п. 5.1, 5.2, 5.3).

**Критерии балльной оценки различных форм текущего контроля**

Текущий контроль заключается в контролеуспеваемости, проводимом на практических занятиях при выполнении лабораторных и контрольной работы:

* проверка знаний, полученных в процессе выполнения самостоятельных индивидуальных практических заданий в компьютерном классе;
* проверка знаний, в результате выполнения внеаудиторных заданий по изучению теоретического материала;
* проверка правильности и защита самостоятельной лабораторной/контрольной работы.

В процессе защиты выполненные **лабораторные** работы оцениваются баллами. Критерии оценки приведены в табл. 7, где знания за все занятия максимально оцениваются 10-ю баллами (указаны через дробь).

Таблица 5

Критерии оценивания в процессе выполнения лабораторных работ

|  |  |  |
| --- | --- | --- |
| **Критерии оценивания компетенций** | | **Оценка в баллах** |
| 1. | Практическое задание выполнено полностью и правильно. В работе продемонстрирована активность и умение находить правильные решения. На защите продемонстрировано знание инструментов, используемых для ее решения. Продемонстрировано уверенное владение всеми изучаемыми инструментальными средствами: MS Excel, LensKit, Azure ML, Python/R (если применимо). Продемонстрировано владение средствами поиска информационных ресурсов в локальных и глобальных сетях. Проявлена активность в интерактивной форме обучения. | 10/10 |
| 3. | Задание выполнено не полностью, допущены ошибки. Обнаружено знание только базовой части программного материала. Допущены неточности в ответах, неполные и/или противоречивых высказывания.  Трудности в использовании информационных технологий для решения финансовых задач. Проблемы с анализом результатов, полученных в процессе решения задач. Задание выполнено с погрешностями и задержками по срокам Отсутствие активности в интерактивной форме обучения. | 5/10 |
| 4. | Задание не выполнено. Продемонстрировано незнание программного материала и неумение пользоваться информационными технологиями. Многократные пропуски занятий без уважительных причин. Не участие в интерактивной форме обучения. | 0/10 |

**Контрольная работа** проводится в виде соревнования на платформе [***kaggle***](https://www.kaggle.com/)***.com.***

Участники соревнования строят рекомендательные системы на основе предоставленных преподавателем данных (данные могут содержать шумы, выбросы, пропущенные значения и требовать дополнительной обработки, а также формирования синтетических признаков). Цель работы – выдача рекомендаций, качество которых оценивается по заранее определенному критерию (метрике качества).

Ссылка на чемпионат выдается участникам преподавателем.

Исходный полный набор данных разделен на три набора:

* набор, предназначенный для обучения модели (информация о рейтингах доступна)
* публичный тестовый набор данных (информация о рейтингах не доступна);
* закрытый тестовый набор данных (информация о рейтингах не доступна).

Участник соревнования должен загрузить на платформу чемпионата файл с прогнозами тестовому набору данных, половина из которых относится к публичному тестовому набору данных, другая половина - к закрытому тестовому набору данных.

Сразу после загрузки результатов автоматически рассчитывается значение метрики качества и рейтинг участников на основании публичного тестового набора данных. В момент окончания соревнования для каждого участника происходит перерасчет доли правильных ответов и рейтинга на основании закрытого тестового набора данных.

Оценка контрольной работы зависит от значения метрики качества и также приводится к 10-бальной шкале.

При расчете итоговой оценки работы в семестре весовой коэффициент, присваиваемый оценке контрольной работы, в 2-4 превышает вес лабораторной работы.

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

## 7.1 перечень компетенций с указанием индикаторов их достижения*[[1]](#footnote-1)* в процессе освоения образовательной программы

Перечень компетенций, формируемых в процессе освоения дисциплины содержится в разделе 2 «Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы».

## 7.2 типовые контрольные задания или иные материалы, необходимые для оценки индикаторов достижения компетенций, умений и знаний;

Теоретические вопросы для подготовки к экзамену определяются содержанием раздела 5.1. Практические задания, выполнение которых необходимо для достижения компетенций, знаний и умений, отражены в тематике лабораторных работ (раздел 5.2).

**Пример задания контрольной/лабораторной работы:**

1. Загрузить данные из источника, указанного преподавателем – например, база данных об отзывах клиентов ресторанов или рейтингах фильмов (доступны в репозитории Azure ML)
2. Исследовать данные с помощью одного или нескольких программных инструментов (Excel/LensKit/MS Azure/Python/R), сформулировать выводы.
3. В соответствии с методологией, разобранной на лекции, а также следуя указаниям преподавателя, построить рекомендательную систему одного из следующих типов: неперсонализированная/слабоперсонализированная, контентно-ориентированная, система на основе коллаборативной фильтрации, гибридная.
4. При построении рекомендательной системы может потребоваться дополнительная обработка данных: устранение шумов, выбросов, пропущенных значений, снижение размерности матрицы рейтингов, а также формирования синтетических признаков (в том числе с применением TFIDF анализа).
5. Оценить качество построенной системы, рассчитав базовые и продвинутые числовые метрики, в соответствии с методологией, разобранной на лекции;
6. Сделать выводы о результатах лабораторной работы, применимости/ целесообразности использования того или иного подхода в зависимости от постановки задачи, предложить направления улучшения результата.

## 7.3 Приказы, распоряжения ректората о контроле уровня освоения дисциплин и сформированности компетенций студентов.

Соответствующие приказы, распоряжения ректората о контроле уровня освоения дисциплин и сформированности компетенций студентов, актуальные на момент проведения контроля.

# 8. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины

**Основная литература:**

1. Дадян Э.Г. Методы, модели, средства хранения и обработки данных: учебник / Э.Г. Дадян, Ю.А. Зеленков. — Москва: Вузовский учебник: ИНФРА-М, 2017. — 168 с. – То же [Электронный ресурс]. - Режим доступа: http://znanium.com/catalog/product/543943.
2. Калинина В.Н. Анализ данных: Компьютерный практикум / В.Н. Калинина, В.И. Соловьев. – Москва: КНОРУС, 2017. – 166 с. – То же [Электронный ресурс]. – Режим доступа: https://www.book.ru/book/929386.

**Дополнительная литература:**

1. Волкова В. Н. Моделирование систем и процессов [Электронный ресурс]: учебник / В. Н. Волкова [и др.]; под ред. В. Н. Волковой, В. Н. Козлова. — Москва: Издательство Юрайт, 2018. — 450 с. – Режим доступа: https://biblio-online.ru/book/E7D370B9-3C64-4A0F-AF1B-F6BD0EEEBCD0/modelirovanie-sistem-i-processov.
2. Волкова В. Н. Моделирование систем и процессов. Практикум [Электронный ресурс]: учебное пособие / В. Н. Волкова [и др.]; под ред. В. Н. Волковой. — Москва: Издательство Юрайт, 2018. — 295 с. - Режим доступа: https://biblio-online.ru/book/3DF77B78-AF0B-48EE-9781-D60364281651/modelirovanie-sistem-i-processov-praktikum.
3. Колдаев В.Д. Структуры и алгоритмы обработки данных [Электронный ресурс]: учебное пособие / В.Д. Колдаев. - Москва: ИЦ РИОР: НИЦ ИНФРА-М, 2014. - 296 с. - Режим доступа: http://znanium.com/catalog/product/418290.

# 9. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины

1. Информационно-образовательный портал Финансового университета при Правительстве Российской Федерации <http://portal.ufrf.ru/>.
2. Сайт кафедры департамента анализа данных, принятия решений и финансовых технологий. <http://fa.ru/dep/data_analysis/>
3. Электронная библиотека Финансового университета (ЭБ) <http://elib.fa.ru/> (<http://librarry.fa.ru/files/elibfa.pdf>)
4. Массовый открытый онлайн-курс/специализация «Microsoft Professional Program in Data Science/ Microsoft.»- <https://www.edx.org/microsoft-professional-program-data-science#edx-product-discovery-cards>
5. Массовый открытый онлайн-курс/специализация «Машинное обучение и анализ данных» / МФТИ и Яндекс. – <https://www.coursera.org/specializations/machine-learning-data-analysis>
6. Массовый открытый онлайн-курс/специализация “Recommender Systems”/ University of Minnesota – https://www.coursera.org/specializations/recommender-systems
7. Массовый открытый онлайн-курс/специализация “Machine Learning”/ Stanford University - https://www.coursera.org/learn/machine-learning/home/welcome
8. Профессиональный ресурс по машинному обучению. - <https://stackoverflow.com>
9. Профессиональный ресурс по машинному обучению.- <https://stackexchange.com>
10. Платформа для соревнований по машинному обучению – [www.kaggle.com](http://www.kaggle.com)
11. Электронно-библиотечная система BOOK.RU http://www.book.ru
12. Электронно-библиотечная система Znanium <http://www.znanium.com>
13. Электронно-библиотечная система издательства «ЮРАЙТ» https://www.biblio-online.ru/

# 10. Методические указания для обучающихся по освоению дисциплины

Методические указания для обучающихся по освоению дисциплины (комплекс рекомендаций и разъяснений, позволяющий студенту оптимальным образом организовать процесс изучения учебного материала дисциплины) представлены в **Учебно-организационном комплексе для дисциплин Департамента анализа данных, принятия решений и финансовых технологий**, размещенном на странице Департамента анализа данных, принятия решений и финансовых технологий сайта Финансового университета.

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

1. Информационно-образовательный портал Финансового университета при Правительстве Российской Федерации <http://portal.ufrf.ru/>.
2. Microsoft Excel
3. Microsoft Azure (Cortana Intelligence Suite)
4. LensKit (требуется поддержка Java)
5. Python/ R

# 12. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине.

Для освоения дисциплины необходимо любое вычислительное средство – компьютер, смартфон или планшет. Предпочтительно использование компьютера под управлением одной из операционных систем *Windows*, *Linux* или *Mac OSX* с доступом к сети «Интернет». Практические занятия должны проводиться в компьютерных классах университета.

1. Указываются при реализации актуализированных ОС ВО ФУ и ФГОС ВО3++ [↑](#footnote-ref-1)