


Федеральное государственное образовательное бюджетное учреждение  
высшего образования  
**«ФИНАНСОВЫЙ УНИВЕРСИТЕТ  
ПРИ ПРАВИТЕЛЬСТВЕ РОССИЙСКОЙ ФЕДЕРАЦИИ»**  
(Финансовый университет)

Департамент анализа данных,  
принятия решений и финансовых технологий

**УТВЕРЖДАЮ**

Проректор по развитию  
образовательных программ и  
международной деятельности  
 Е.А. Каменева

« 28 » ноября 2017 г.

**В.И. Соловьев, В.Г. Феклин**

**ПРИКЛАДНЫЕ ЗАДАЧИ  
МАШИННОГО ОБУЧЕНИЯ**

Рабочая программа дисциплины  
для студентов, обучающихся  
по направлениям подготовки 01.03.02 «Прикладная математика и информатика», 09.03.03 «Прикладная информатика», 38.03.04 «Государственное и муниципальное управление», 38.03.05 «Бизнес-информатика», 39.03.01 «Социология», 41.03.04 «Политология», 42.03.01 «Реклама и связи с общественностью», 38.03.01 «Экономика», 38.03.02 «Менеджмент», 38.03.03 «Управление персоналом», 40.03.01 «Юриспруденция», 10.03.01 «Информационная безопасность», 43.03.02 «Туризм»

*Рекомендовано Ученым советом факультета  
«Прикладная математика и информационные технологии»  
(протокол № 48 от 21 ноября 2017 г.)  
Одобрено Департаментом анализа данных, принятия решений и  
финансовых технологий  
(протокол № 4 от 21 ноября 2017 г.)*

**Москва – 2017**

**Федеральное государственное образовательное бюджетное учреждение  
высшего образования  
«ФИНАНСОВЫЙ УНИВЕРСИТЕТ  
ПРИ ПРАВИТЕЛЬСТВЕ РОССИЙСКОЙ ФЕДЕРАЦИИ»  
(Финансовый университет)**

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

**В.И. Соловьев, В.Г. Феклин**

**ПРИКЛАДНЫЕ ЗАДАЧИ  
МАШИННОГО ОБУЧЕНИЯ**

**Рабочая программа дисциплины**

для студентов, обучающихся  
по направлениям подготовки 01.03.02 «Прикладная математика и  
информатика», 09.03.03 «Прикладная информатика», 38.03.04  
«Государственное и муниципальное управление», 38.03.05 «Бизнес-  
информатика»,  
39.03.01 «Социология», 41.03.04 «Политология», 42.03.01 «Реклама и связи  
с общественностью», 38.03.01 «Экономика», 38.03.02 «Менеджмент»,  
38.03.03 «Управление персоналом», 40.03.01 «Юриспруденция»,  
10.03.01 «Информационная безопасность», 43.03.02 «Туризм»

**Москва – 2017**

**Федеральное государственное образовательное бюджетное учреждение  
высшего образования**

**«ФИНАНСОВЫЙ УНИВЕРСИТЕТ  
ПРИ ПРАВИТЕЛЬСТВЕ РОССИЙСКОЙ ФЕДЕРАЦИИ»  
(Финансовый университет)**

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

**УТВЕРЖДАЮ**

Проректор по развитию  
образовательных программ и  
международной деятельности

\_\_\_\_\_ Е.А. Каменева

28.11.2017 г.

**В.И. Соловьев, В.Г. Феклин**

## **ПРИКЛАДНЫЕ ЗАДАЧИ МАШИННОГО ОБУЧЕНИЯ**

Рабочая программа дисциплины  
для студентов, обучающихся  
по направлениям подготовки 01.03.02 «Прикладная математика и  
информатика», 09.03.03 «Прикладная информатика», 38.03.04  
«Государственное и муниципальное управление», 38.03.05 «Бизнес-  
информатика»,  
39.03.01 «Социология», 41.03.04 «Политология», 42.03.01 «Реклама и связи  
с общественностью», 38.03.01 «Экономика», 38.03.02 «Менеджмент»,  
38.03.03 «Управление персоналом», 40.03.01 «Юриспруденция»,  
10.03.01 «Информационная безопасность», 43.03.02 «Туризм»

*Рекомендовано Ученым советом факультета  
«Прикладная математика и информационные технологии»  
(протокол № 48 от 21 ноября 2017 г.)  
Одобрено Департаментом анализа данных, принятия решений и  
финансовых технологий  
(протокол № 4 от 21 ноября 2017 г.)*

## СОДЕРЖАНИЕ

1. Наименование дисциплины .....	4
2. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы .....	4
3. Место дисциплины в структуре основной образовательной программы	6
4. Объем дисциплины и виды учебной работы .....	7
5. Содержание дисциплины, структурированное по темам .....	8
5.1. Содержание тем дисциплины .....	8
5.2. Учебно-тематический план .....	11
5.3. Содержание практических занятий .....	16
6. Перечень основной и дополнительной учебной литературы .....	17
7. Перечень ресурсов информационно-телекоммуникационной сети «Интернет» .	
8. Методические указания для обучающихся по освоению дисциплины	17
9. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине .....	19
10. Материально-техническая база, необходимая для осуществления образовательного процесса по дисциплине .....	19

## **1.Наименование дисциплины**

Дисциплина «Прикладные задачи машинного обучения».

Основная цель данной дисциплины – формирование у студентов представления об основных методах машинного обучения и прикладных задачах, решаемых с помощью этих методов.

## **2 . Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы**

Дисциплина «Прикладные методы машинного обучения» обеспечивает инструментарий формирования следующих компетенций.

Направления подготовки «Прикладная математика и информатика», «Прикладная информатика», «Государственное и муниципальное управление», «Бизнес-информатика», «Социология», «Юриспруденция» «Политология», «Реклама и связи с общественностью».

**Компетенция ОК-1** – Способность использовать основы философских знаний для формирования мировоззренческой позиции.

*Знать* основные методы машинного обучения.

*Уметь* строить модели классификации, регрессии, кластеризации и поиска аномалий.

*Владеть* профессиональной терминологией в области машинного обучения.

**Компетенция ОК-7** – Способность к самоорганизации и самообразованию. *Знать* основные классы практических задач в области экономики, управления и финансов, решаемых методами машинного обучения.

*Уметь* применять модели машинного обучения для решения практических задач в области предоставления финансовых услуг.

*Владеть* навыками анализа и обработки финансовой информации.

Направления подготовки «Экономика», «Менеджмент»

**Компетенция ОНК-1** Способность использовать основы философских знаний и

основные научные законы для формирования мировоззренческой позиции и в профессиональной деятельности

*Знать* основные классы практических задач в области экономики, управления и финансов, решаемых методами машинного обучения.

*Уметь* применять модели машинного обучения для решения практических задач в области предоставления финансовых услуг.

*Владеть* навыками анализа и обработки финансовой информации.

**Компетенция ОНК-2** Владение культурой мышления, способность анализировать этапы и закономерности исторического развития общества для формирования гражданской позиции

*Знать* основные методы машинного обучения.

*Уметь* строить модели классификации, регрессии, кластеризации и поиска аномалий.

*Владеть* профессиональной терминологией в области машинного обучения.

Направление подготовки «Управление персоналом».

**Компетенция ОНК-1** - Способность использовать основные научные законы в профессиональной деятельности

*Знать* основные классы практических задач в области экономики, управления и финансов, решаемых методами машинного обучения.

*Уметь* применять модели машинного обучения для решения практических задач в области предоставления финансовых услуг.

*Владеть* навыками анализа и обработки финансовой информации.

**Компетенция ОНК-2** - Владение культурой мышления, способность к восприятию, анализу и мировоззренческой оценке происходящих процессов и закономерностей.

*Знать* основные методы машинного обучения.

*Уметь* строить модели классификации, регрессии, кластеризации и поиска аномалий.

*Владеть профессиональной терминологией в области машинного обучения.*

Направление подготовки «Информационная безопасность».

**Компетенция ОК-1** – Способность использовать основы философских знаний для формирования мировоззренческой позиции.

*Знать* основные методы машинного обучения.

*Уметь* строить модели классификации, регрессии, кластеризации и поиска аномалий.

*Владеть* профессиональной терминологией в области машинного обучения.

**Компетенция ОК-8** – Способность к самоорганизации и самообразованию. *Знать* основные классы практических задач в области экономики, управления и финансов, решаемых методами машинного обучения.

*Уметь* применять модели машинного обучения для решения практических задач в области предоставления финансовых услуг.

*Владеть* навыками анализа и обработки финансовой информации.

Направление подготовки «Туризм».

**Компетенция ОК - 1** – Способность использовать основы философских знаний, анализировать главные этапы и закономерности исторического развития для создания социальной значимости своей деятельности.

*Знать* основные методы машинного обучения.

*Уметь* строить модели классификации, регрессии, кластеризации и поиска аномалий.

*Владеть* профессиональной терминологией в области машинного обучения.

**Компетенция ОК - 5** – Способность к самоорганизации и самообразованию.

*Знать* основные классы практических задач в области экономики, управления и финансов, решаемых методами машинного обучения.

*Уметь* применять модели машинного обучения для решения практических задач в области предоставления финансовых услуг.

*Владеть* навыками анализа и обработки финансовой информации.

### **3 .Место дисциплины в структуре основной образовательной программы**

Дисциплина «Прикладные задачи машинного обучения» является дисциплиной по выбору студента Университетского блока.

В процессе изучения дисциплины происходит знакомство с основными понятиями и методами машинного обучения. Студенты приобретают опыт практического использования изучаемых методов в практических задачах, связанных с финансовыми технологиями.

#### **4. Объем дисциплины и виды учебной работы**

<b>Вид учебной работы по дисциплине</b>	<b>Всего (в з/е и часах)</b>	<b>Семестр 3 или 4 или 5</b>
<b>Общая трудоёмкость дисциплины</b>	<b>3/108</b>	<b>3/108</b>
<b>Контактная работа - Аудиторные занятия</b>	<b>36</b>	<b>36</b>
<i>Лекции</i>	<b>18</b>	<b>18</b>
<i>Практические и семинарские занятия,</i>	<b>18</b>	<b>18</b>
<i>в т.ч. занятия в интерактивных формах</i>	<b>16</b>	<b>16</b>
<b>Самостоятельная работа</b>	<b>72</b>	<b>72</b>
Вид текущего контроля	-	-
Вид промежуточной аттестации	<b>Зачет</b>	<b>Зачет</b>

#### **5. Содержание дисциплины, структурированное по темам**

##### **5.1. Содержание тем дисциплины**

##### **Тема 1. От данных к ценности: введение в науки о данных**

Основные определения: интеллектуальный анализ данных, большие данные, машинное обучение.

Методы и задачи машинного обучения, области применения методов и технологий машинного обучения.

Примеры задач машинного обучения: поиск информации в интернете,



распознавание изображений, лиц, эмоций, пола, возраста, ..., распознавание речи, языка, эмоциональной окраски текстов, прогнозирование продаж, прогнозирование оттока клиентов, кредитный скоринг, рекомендательные системы и др.

Цикл обработки данных: поиск данных, сбор данных, очистка данных, трансформация данных, интеллектуальный анализ данных, интерпретация и практическое применение результатов.

Статистические основы обработки данных.

Обзор современных технологий машинного обучения: Microsoft Azure ML, Python, R.

## **Тема 2. Практическое использование моделей классификации**

Принципы разработки и оценки систем машинного обучения. Основные классы моделей машинного обучения: обучение с учителем (классификация и регрессия) и без учителя (кластеризация и поиск аномалий).

Задача классификация с обучением. Задача кредитного скоринга.

Модель логистической регрессии и ее компьютерная реализация.

Методы оценки качества моделей классификации: доля правильных ответов, точность, полнота, F1, AUC. Цены ошибок первого и второго рода.

Модели деревьев и лесов решений. Усиление (бустинг) деревьев решений. Нейронные сети. Компьютерная реализация методов классификации.

Недообучение и переобучение в моделях классификации.

Основы технологии улучшения моделей машинного обучения. Подбор оптимальных параметров моделей. Важность подготовки данных. Генерация синтетических признаков. Работа с пропущенными данными. Работа с несбалансированными выборками.

Лабораторная работа 1. Построение системы кредитного скоринга.

## **Тема 3. Практическое использование моделей регрессии**

Задача прогнозирования продаж. Модель множественной линейной регрессии. Методы оценки качества моделей регрессии. Коэффициент

детерминации, средняя абсолютная ошибка предсказания, средняя относительная ошибка предсказания.

Методы классификации и их компьютерная реализация.

Недообучение и переобучение в моделях регрессии.

Основы технологии улучшения моделей машинного обучения в задачах регрессии.

Лабораторная работа 2. Построение системы прогнозирования продаж.

#### **Тема 4. Практическое использование моделей кластерного анализа и поиска аномалий**

Кластерный анализ и его компьютерная реализация. Задача сегментирования потребителей. Иерархические агломеративные алгоритмы. Метод К-средних.

Методы оценки качества моделей кластерного анализа. Расстояния между кластерами, расстояния между объектами внутри кластеров. Характеристики центров кластеров.

Задача снижения размерности факторного пространства. Метод главных компонент и его компьютерная реализация. Геометрическое и экономическое содержание метода главных компонент. Использование методов снижения размерности для улучшения качества моделей машинного обучения. Использование методов снижения размерности для выявления латентных факторов.

Поиск аномалий. Задача поиска мошеннических транзакций.

Лабораторная работа 3. Сегментирование потребителей.

#### **Тема 5. Построение рекомендательных систем**

Понятие рекомендательной системы, типы оценок. Примеры рекомендательных сервисов.

Принципы создания рекомендательных систем: коллаборативная фильтрация, контентная фильтрация, гибридные подходы.

Алгоритмы, используемые рекомендательными системами: корреляция Пирсона, алгоритмы кластеризации, байесовские сети доверия, цепи Маркова, метод Роккио.

Проблемы рекомендательных систем.

Компьютерная реализация рекомендательных систем.

Лабораторная работа 4. Построение рекомендательного сервиса.

## 5.2. Учебно-тематический план

№ п/п	Наименование темы (раздела) дисциплины	Трудоёмкость в часах						Формы текущего контроля успеваемости
		Всего часов	Аудиторная работа				Самостоятельная работа	
			Общая	Лекции	Практ. и семинарские занятия	Занятия в интерактивных формах		
1	От данных к ценности: Введение в науки о данных	12	4	2	2	0	8	Практические лабораторные работы (анализ конкретной ситуации и обсуждение результатов)
2	Практическое использование моделей классификации	24	8	4	4	4	16	
3	Практическое использование моделей регрессии	24	8	4	4	4	16	
4	Практическое использование моделей кластерного анализа и поиска аномалий	24	8	4	4	4	16	
5	Построение рекомендательных систем	24	8	4	4	4	16	
	Итого	108	36	18	18	16/44 %	72	

### **5.3. Содержание практических занятий**

Практические занятия направлены на интенсивное овладение практическими навыками построения систем машинного обучения, что отражено в форме их проведения: лабораторные работы.

#### **1. Знакомство с Microsoft Azure ML. Работа с данными**

Получение бесплатной подписки на базовый уровень Microsoft Azure ML. Знакомство с интерфейсом Microsoft Azure ML. Подготовка и загрузка набора данных. Создание эксперимента и проекта.

Рекомендуемые источники: [2, 10].

#### **2-3 Построение системы кредитного скоринга**

Лабораторная работа «Построение системы кредитного скоринга» (построение компьютерной системы машинного обучения, анализ конкретной ситуации и обсуждение результатов).

Интерактив: построение компьютерной системы машинного обучения, анализ конкретной ситуации и обсуждение результатов – 100% от трудоемкости семинарского занятия.

Рекомендуемые источники: [1, 2, 3, 4].

#### **4-5 Построение системы прогнозирования продаж**

Лабораторная работа «Построение системы прогнозирования продаж» (построение компьютерной системы машинного обучения, анализ конкретной ситуации и обсуждение результатов).

Интерактив: построение компьютерной системы машинного обучения, анализ конкретной ситуации и обсуждение результатов – 100% от трудоемкости семинарского занятия.

Рекомендуемые источники: [1, 2, 3, 4].

## **6-7 Сегментирование потребителей**

Лабораторная работа «Сегментирование потребителей» (построение компьютерной системы машинного обучения, анализ конкретной ситуации и обсуждение результатов).

Интерактив: построение компьютерной системы машинного обучения, анализ конкретной ситуации и обсуждение результатов – 100% от трудоемкости семинарского занятия.

Рекомендуемые источники: [1, 2, 4].

## **8-9 Построение рекомендательного сервиса**

Лабораторная работа «Построение рекомендательного сервиса» (построение компьютерной системы машинного обучения, анализ конкретной ситуации и обсуждение результатов).

Интерактив: построение компьютерной системы машинного обучения, анализ конкретной ситуации и обсуждение результатов – 100% от трудоемкости семинарского занятия.

Рекомендуемые источники: [1, 2, 3, 4].

## **6.Перечень основной и дополнительной учебной литературы**

### **а) основная:**

1. Миркин, Б.Г. Введение в анализ данных [Электронный ресурс]: учебник и практикум / Б. Г. Миркин. — Москва: Юрайт, 2018. — 174 с. — Режим доступа: <https://biblio-online.ru/book/46A41F93-BC46-401C-A30E-27C0FB60B9DE>

2. Калинина, В.Н. Анализ данных: компьютерный практикум / В.Н. Калинина, В.И. Соловьев. — Москва: КНОРУС, 2017. — 166 с.

### **б) дополнительная**

1. Воронцов, К.В. Машинное обучение [Электронный ресурс]: курс лекций / К.В. Воронцов. — Режим доступа: [http://www.machinelearning.ru/wiki/index.php?title=Машинное\\_обучение\\_\(курс](http://www.machinelearning.ru/wiki/index.php?title=Машинное_обучение_(курс))

## **7.Перечень ресурсов информационно-телекоммуникационной сети «Интернет»**

4.Информационно-образовательный портал Финансового университета при Правительстве Российской Федерации <http://portal.ufrf.ru/>.

5 .Сайт кафедры департамента анализа данных, принятия решений и финансовых технологий. [http://fa.ru/dep/data\\_analysis/](http://fa.ru/dep/data_analysis/)

6 .Документация по Azure ML  
<https://docs.microsoft.com/ru-ru/azure/machine-learning/studio/>

7 .*Applied Machine Learning* / Microsoft. –  
<https://www.edx.org/course/applied-machine-learning-microsoft-dat203-3x-1>

8 .*Data Science Essentials* / Microsoft. – <https://www.edx.org/course/data-science-essentials-microsoft-dat203-1x-3>

9 .*Principles of Machine Learning* / Microsoft. –  
<https://www.edx.org/course/principles-machine-learning-microsoft-dat203-2x-3>

10 .Профессиональный ресурс по машинному обучению. –  
<https://machinelearning.ru>

11 .Специализация «Машинное обучение и анализ данных» / МФТИ и Яндекс. – <https://www.coursera.org/specializations/machine-learning-data-analysis>

12 .Электронная библиотека Финансового университета (ЭБ) <http://elib.fa.ru/>  
(<http://library.fa.ru/files/elibfa.pdf>)

13 .Электронно-библиотечная система издательства «ЮРАЙТ»  
<https://www.biblio-online.ru/>

## **8.Методические указания для обучающихся по освоению дисциплины**

Самостоятельная работа студентов проходит аудиторно и внеаудиторно. Организации самостоятельной работы служит календарнотематический план изучения дисциплины. В этом плане указана тематика лекций и практических занятий.

При подготовке к лекции целесообразно предварительно познакомиться с ее содержанием по рекомендованным источникам и выделить наиболее трудные вопросы. Во время лекций следует конспектировать содержание лекции. После занятий следует провести работу с конспектом: отредактировать записи, оформить конспект. При оформлении целесообразно выделять специальным образом названия тем и формулировки вопросов, основные определения, формулировки теорем и примеры. Сделанные записи нужно сопоставить с учебниками и учебными пособиями и в случае расхождений проконсультироваться с преподавателем.

### **Методические указания по проведению практических занятий**

*Практические занятия* структурно состоят из следующих компонент:

1. проверка наличия выполненного задания самостоятельной работы каждого студента;
2. выборочная проверка корректности выполнения домашнего задания;
3. разбор типичных ошибок, возникших в самостоятельной работе;
4. рассмотрение теоретических вопросов, связанных с текущим практическим занятием;
5. разбор методов выполнения практических заданий и решения задач;
6. корректировка заданий для самостоятельной работы студентов;
7. интерактивная форма – решение лабораторных работ по тематике занятия в малых группах (2-4 студента).

При подготовке к практическому занятию необходимо повторить или, если это требуется, изучить соответствующий теоретический материал. Во время занятия нужно точно следовать указаниям преподавателя к решению лабораторных работ. После занятий необходимо просмотреть полученные решения и восстановить имеющиеся пробелы или закончить решение. В случае затруднений отметить соответствующие задания и обратиться за консультацией к преподавателю.

Практические занятия проходят в интерактивной форме. Наряду с



правильностью решения преподаватель учитывает также и активность студентов, направленную на решение предложенных задач.

## **9.Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине**

9.1. Комплект лицензионного программного обеспечения:

Windows Microsoft Office,

Антивирус ESET Endpoint Security

9.2. Современные профессиональные базы данных и информационные справочные системы:

Справочная правовая система «КонсультантПлюс» (<http://www.consultant.ru>).

Справочная правовая система «Гарант» (<http://www.garant.ru>).

Информационно-образовательный портал Финансового университета. - <http://portal.ufrf.ru>.

Azure ML

R и RStudio

Anaconda и Python

9.3. Сертифицированные программные и аппаратные средства защиты информации – не предусмотрено.

## **10.Материально-техническая база, необходимая для осуществления образовательного процесса по дисциплине**

Для освоения дисциплины необходимо любое вычислительное средство – компьютер, смартфон или планшет. Предпочтительно использование компьютера. При этом возможно использование компьютеров компьютерных классах университета.

Все изучаемые технологии доступны на личных устройствах студентов в любой точке, где доступна сеть Интернет.