

МИНОБРНАУКИ РОССИИ
ФГБОУ ВО «БАШКИРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»
ИНСТИТУТ ЭКОНОМИКИ, ФИНАНСОВ И БИЗНЕСА

ИНСТИТУТ ЭКОНОМИКИ, ФИНАНСОВ И БИЗНЕСА

Утверждено:
на заседании кафедры
протокол от «17» февраля 2021 г. №5

Зав. кафедрой  /Р.Х.Бахитова

Согласовано:
Председатель УМК института

 /Л.Р. Абзалилова

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Методы искусственного интеллекта в экономике

Часть, формируемая участниками образовательных отношений

Программа магистратуры

Направление подготовки
38.04.05 Бизнес-информатика

Направленность (профиль) подготовки
«Информационная бизнес-аналитика»

Квалификация
магистр

Разработчик (составитель) РПД:
доцент, к-т техн..наук



Лакман И.А.

Для приема 2021 г.

Уфа 2021 г.

Составитель / составители: к-т техн. наук, доцент Лакман И.А.

Рабочая программа дисциплины утверждена на заседании кафедры цифровой экономики и коммуникации протокол от «17» февраля 2021 г. №5

Дополнения и изменения, внесенные в рабочую программу дисциплины, утверждены на заседании кафедры _____,
протокол № ____ от « ____ » _____ 20 _ г.

Заведующий кафедрой _____ / _____ Ф.И.О/

Дополнения и изменения, внесенные в рабочую программу дисциплины, утверждены на заседании кафедры _____,
протокол № ____ от « ____ » _____ 20 _ г.

Заведующий кафедрой _____ / _____ Ф.И.О/

Список документов и материалов

1. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы	4
2. Место дисциплины в структуре ООП	6
3. Содержание рабочей программы (объем дисциплины, типы и виды учебных занятий, учебно-методическое обеспечение самостоятельной работы обучающихся).....	6
4. Фонд оценочных средств по дисциплине	11
4.1 Перечень компетенций и индикаторов достижения компетенций с указанием соотнесенных с ними запланированных результатов обучения по дисциплине. Описание критериев и шкал оценивания результатов обучения по дисциплине	11
4.2. Типовые контрольные задания или иные материалы, необходимые для оценивания результатов обучения по дисциплине, соотнесенных с установленными в образовательной программе индикаторами достижения компетенций. Методические материалы, определяющие процедуры оценивания результатов обучения по дисциплине	13
5. Учебно-методическое и информационное обеспечение дисциплины	31
5.1. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины.....	31
5.2. Перечень ресурсов информационно-телекоммуникационной сети «Интернет» и программного обеспечения, необходимых для освоения дисциплины	34
6. Материально-техническая база, необходимая для осуществления образовательного процесса по дисциплине	35

1. Перечень планируемых результатов обучения по дисциплине, соотнесённых с планируемыми результатами освоения образовательной программы

По итогам освоения дисциплины **Методы искусственного интеллекта в экономике** обучающийся должен достичь следующих результатов обучения:

Категория (группа) компетенций ¹ (при наличии ОПК)	Формируемая компетенция (с указанием кода)	Код и наименование индикатора достижения компетенции	Результаты обучения по дисциплине
-	<i>ПК-1 Способен анализировать и выбирать оптимальные подходы к реализации процедур бизнес-анализа в организации</i>	<i>ПК 1.1 Применяет подходы к проведению бизнес-анализа, работе с информацией и требованиями, оценке эффективности работы по бизнес-анализу и умеет обосновать их выбор</i>	<p><i>Знать:</i> основные принципы использования машинного обучения для задач кредитного риск-менеджмента (PD, EAD, LGD) основные принципы использования моделей выживаемости для задач коллекторского скоринга.</p> <p><i>Уметь:</i> применять инструменты машинного обучения для задач кредитного риск-менеджмента (PD, EAD, LGD), для задач кредитного скоринга, для задач прогноза оттока клиентов, для задач выявления мошеннических транзакций</p> <p><i>Владеть:</i> методами применения алгоритмов машинного обучения в риск-менеджменте с использованием современных информационных средств и сред</p>
		<i>ПК 1.2 Определяет связи и зависимости между элементами информации бизнес-анализа</i>	<p><i>Знать:</i> основные методы нейросетевого моделирования и глубокого обучения</p> <p><i>Уметь:</i> методы нейросетевого моделирования и глубокого обучения для решения задач в экономике и финансах</p> <p><i>Владеть:</i> нейросетевого моделирования и глубокого обучения для решения задач в экономике и финансах в современных средах</p>

¹ Указывается только для УК и ОПК (при наличии).

2. Место дисциплины в структуре ООП

Дисциплина «**Методы искусственного интеллекта в экономике**» является дисциплиной по выбору и относится к части, формируемой участниками образовательных отношений.

Дисциплина изучается на 2 курсе магистратуры в 4 семестре и на 3-ем курсе в 5-ом семестре.

Целью изучения дисциплины «**Методы искусственного интеллекта в экономике**» является формирование знаний, навыков и умений, необходимых для понимания и практической реализации современных концепций интеллектуального анализа текста, относящегося к неструктурированной информации.

Для успешного освоения курса необходимы компетенции, сформированные в рамках курса бакалавриата теории вероятностей и математической статистики, а также магистерского курса **Анализа данных**.

3. Содержание рабочей программы (объем дисциплины, типы и виды учебных занятий, учебно-методическое обеспечение самостоятельной работы обучающихся)

МИНОБРНАУКИ РОССИИ
ФГБОУ ВО «БАШКИРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»
ИНСТИТУТ ЭКОНОМИКИ, ФИНАНСОВ И БИЗНЕСА

СОДЕРЖАНИЕ РАБОЧЕЙ ПРОГРАММЫ

дисциплины «Методы искусственного интеллекта в экономике»
на 4 семестр
очно-заочной формы обучения

Вид работы	Объем дисциплины
Общая трудоемкость дисциплины (ЗЕТ / часов)	5/180
Учебных часов на контактную работу с преподавателем:	
лекций	8
практических/ семинарских	16
лабораторных	
других (групповая, индивидуальная консультация и иные виды учебной деятельности, предусматривающие работу обучающихся с преподавателем) (ФКР)	0,2
Учебных часов на самостоятельную работу обучающихся (СР)	155,8
Учебных часов на подготовку к экзамену/зачету/дифференцированному зачету (Контроль)	

Форма(ы) контроля:
Зачет 4 семестр

№ п/п	Тема и содержание	Форма изучения материалов: лекции, практические занятия, семинарские занятия, лабораторные работы, самостоятельная работа и трудоемкость (в часах)					Основная и дополнительная литература, рекомендуемая магистрантам (номера из списка)	Задания по самостоятельной работе магистрантов	Форма текущего контроля успеваемости (коллоквиумы, контрольные работы, компьютерные тесты и т.п.)
		Всего	ЛК	ПР/СЕМ	ЛР	СРС			
1	2	3	4	5	6	7	8	9	10
Модуль 1. Методы машинного обучения для кредитного риск-менеджмента									
1.	EAD-моделирование на основе алгоритмов машинного обучения (RF) для определения кредитных обязательств по договору на момент дефолта,	30	2	4		24	№№ 4, 5, из основного списка, №№ 1 из дополнительного списка	Работа с литературой и другими рекомендуемыми источниками. Выполнение кейса 1	Оценка кейсов 1
2.	PD-моделирование на основе алгоритмов машинного обучения (CART) для оценки вероятности дефолта клиента	30	2	2		26	№№ 4, 5, из основного списка, №№ 1 из дополнительного списка	Работа с литературой и другими рекомендуемыми источниками. Выполнение кейса 2	Оценка кейсов 2
3.	LGD-моделирование на основе алгоритмов машинного обучения (RF) для определения уровня потерь при дефолте	30	2	4		24	№№ 1,2 из основного списка, №№ 3, 4 из дополнительного списка	Работа с литературой и другими рекомендуемыми источниками. Выполнение кейса 3	Проверка выполнения кейса 3
4.	•Кредитный скоринг на основе логистической регрессии, критерий Хосмер-Лемешова.	30	2	2		26	№№ 1,2 из основного списка,	Работа с литературой и другими рекомендуемыми	Проверка выполнения кейса 4

	Маржинальные эффекты. Стресс-тестирование.						№№ 3, 4 из дополнительного списка	источниками. Выполнение кейса 4	
5.	Итоговый контроль по Модулю 1	7,8				7,8		Подготовка к устному опросу	Устный опросу
Модуль 2. Задачи коллекторского скоринга на основе машинного обучения									
7.	Задача коллекторского скоринга на основе моделей выживаемости. Кривые Каплана-Майера. Выявление предикторов риска разорения на основе модели Кокса	24	1	2		21	№№ 2, 3 из основного списка, №№ 1, 4 из дополнительного списка	Работа с литературой и другими рекомендуемыми источниками. Выполнение кейса 5	Проверка выполнения кейса 5
8.	•задача прогноза оттока клиентов на основе модели выживаемости (модели ускоренной жизни AFT)	24	1	2		21	№№ 2, 3 из основного списка, №№ 1, 4 из дополнительного списка	Работа с литературой и другими рекомендуемыми источниками. Выполнение кейса 5 (часть 2)	Проверка выполнения кейса 5
9.	Итоговый контроль по модулю 3	4				4		Подготовка к устному опросу	Устный опросу
10.	ФКР	0,2				0,2			
11.	Зачет								

МИНОБРНАУКИ РОССИИ
ФГБОУ ВО «БАШКИРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»
ИНСТИТУТ ЭКОНОМИКИ, ФИНАНСОВ И БИЗНЕСА

СОДЕРЖАНИЕ РАБОЧЕЙ ПРОГРАММЫ

дисциплины «Методы искусственного интеллекта в экономике»
на 5 семестр
очно-заочной формы обучения

Вид работы	Объем дисциплины
Общая трудоемкость дисциплины (ЗЕТ / часов)	
	2/72
Учебных часов на контактную работу с преподавателем:	
лекций	10
практических/ семинарских	18
лабораторных	
других (групповая, индивидуальная консультация и иные виды учебной деятельности, предусматривающие работу обучающихся с преподавателем) (ФКР)	1,2
Учебных часов на самостоятельную работу обучающихся (СР)	6,8
Учебных часов на подготовку к экзамену/зачету/дифференцированному зачету (Контроль)	36

Форма(ы) контроля:

Экзамен 5 семестр

№ п/п	Тема и содержание	Форма изучения материалов: лекции, практические занятия, семинарские занятия, лабораторные работы, самостоятельная работа и трудоемкость (в часах)					Основная и дополнительная литература, рекомендуемая магистрантам (номера из списка)	Задания по самостоятельной работе магистрантов	Форма текущего контроля успеваемости (коллоквиумы, контрольные работы, компьютерные тесты и т.п.)
		Всего	ЛК	ПР/СЕМ	ЛР	СРС			
1	2	3	4	5	6	7	8	9	10
Модуль 3. Методы искусственного интеллекта									
12.	Нейросетевое моделирование в экономике <ul style="list-style-type: none"> • Задача адаптивной фильтрации; Методы безусловной оптимизации; • Однослойный персептрон. • SVM (машина опорных векторов). Ядро – RBF – радиальная базисная функция для SVM • Применение алгоритма SVM для решения задач определения мошеннических транзакций 	16	4	10		2	№№ 2, 4 из основного списка, №№ 1, 4 из дополнительного списка	Работа с литературой и другими рекомендуемыми источниками. Выполнение кейса б	Проверка выполнения кейса б
13.	DeepLearning в экономике: <ul style="list-style-type: none"> • структуры глубоких нейронных сетей • структуры глубоких нейронных сетей • Обратное распространение ошибки 	16	6	8		2	№№ 2,3 из основного списка, №№ 4 из дополнительного списка	Работа с литературой и другими рекомендуемыми источниками. Проверка докладов	Проверка докладов
14.	Итоговый контроль по Модулю 3	2,8				2,8		Подготовка к устному опросу	Устный опрос
15.	ФКР	1,2				1,2			
16.	Экзамен	36				36			
17.	Всего часов:	252	18	34	-	200			

4. Фонд оценочных средств по дисциплине

4.1 Перечень компетенций и индикаторов достижения компетенций с указанием соотнесенных с ними запланированных результатов обучения по дисциплине. Описание критериев и шкал оценивания результатов обучения по дисциплине

Код и формулировка компетенции ПК-1 Способен анализировать и выбирать оптимальные подходы к реализации процедур бизнес-анализа в организации

Код и наименование индикатора достижения компетенции	Результаты обучения по дисциплине	Критерии оценивания результатов обучения			
		2 («Не удовлетворительно»)	3 («Удовлетворительно»)	4 («Хорошо»)	5 («Отлично»)
<i>ПК 1.1 Применяет подходы к проведению бизнес-анализа, работе с информацией и требованиями, оценке эффективности работы по бизнес-анализу и умеет обосновать их выбор</i>	Знать: основные принципы использования машинного обучения для задач кредитного риск-менеджмента (PD, EAD, LGD) основные принципы использования моделей выживаемости для задач коллекторского скоринга	Фрагментарные представления об основных принципах использования машинного обучения для задач кредитного риск-менеджмента (PD, EAD, LGD) основные принципы использования моделей выживаемости для задач коллекторского скоринга	Неполные представления об основных принципах использования машинного обучения для задач кредитного риск-менеджмента (PD, EAD, LGD) основные принципы использования моделей выживаемости для задач коллекторского скоринга	Сформированные, но содержащие отдельные пробелы представления об основных принципах использования машинного обучения для задач кредитного риск-менеджмента (PD, EAD, LGD) основные принципы использования моделей выживаемости для задач коллекторского скоринга	Сформированные систематические представления об основных принципах использования машинного обучения для задач кредитного риск-менеджмента (PD, EAD, LGD) основные принципы использования моделей выживаемости для задач коллекторского скоринга
	Уметь: Применять инструменты машинного обучения для задач кредитного риск-менеджмента (PD, EAD, LGD), для задач кредитного скоринга, для задач прогноза оттока клиентов, для задач выявления	Фрагментарные умения Применять инструменты машинного обучения для задач кредитного риск-менеджмента (PD, EAD, LGD), для задач кредитного скоринга, для задач прогноза оттока клиентов, для задач выявления мошеннических транзакций	В целом успешное, но не систематическое использование умений применять инструменты машинного обучения для задач кредитного риск-менеджмента (PD, EAD, LGD), для задач кредитного скоринга, для задач прогноза оттока	В целом успешное, но содержащее отдельные пробелы использование умений применять инструменты машинного обучения для задач кредитного риск-менеджмента (PD, EAD, LGD), для задач кредитного скоринга, для задач прогноза	Сформированное умение применять инструменты машинного обучения для задач кредитного риск-менеджмента (PD, EAD, LGD), для задач кредитного скоринга, для задач прогноза оттока клиентов, для задач

	мошеннических транзакций		клиентов, для задач выявления мошеннических транзакций	оттока клиентов, для задач выявления мошеннических транзакций	выявления мошеннических транзакций ;
	Владеть: Методами применения алгоритмов машинного обучения в риск-менеджменте с использованием современных информационных средств и сред	Фрагментарное владение методами применения алгоритмов машинного обучения в риск-менеджменте с использованием современных информационных средств и сред	В целом успешное, но не систематическое применение методов применения алгоритмов машинного обучения в риск-менеджменте с использованием современных информационных средств и сред	В целом успешное, но содержащее отдельные пробелы применение методов применения алгоритмов машинного обучения в риск-менеджменте с использованием современных информационных средств и сред	Успешное и систематическое применение методов применения алгоритмов машинного обучения в риск-менеджменте с использованием современных информационных средств и сред
<i>ПК 1.2 Определяет связи и зависимости между элементами информации бизнес-анализа</i>	Знать: основные методы нейросетевого моделирования и глубокого обучения	Фрагментарные представления об основных методах нейросетевого моделирования и глубокого обучения	Неполные представления об основных основные методы нейросетевого моделирования и глубокого обучения	Сформированные, но содержащие отдельные пробелы представления об основных основные методы нейросетевого моделирования и глубокого обучения	Сформированные систематические представления об основных основные методы нейросетевого моделирования и глубокого обучения
	Уметь: Применять методы нейросетевого моделирования и глубокого обучения для решения задач в экономике и финансах	Фрагментарные умения применять методы нейросетевого моделирования и глубокого обучения для решения задач в экономике и финансах	В целом успешное, но не систематическое умение применять методы нейросетевого моделирования и глубокого обучения для решения задач в экономике и финансах	В целом успешное, но содержащее отдельные пробелы в умении применять методы нейросетевого моделирования и глубокого обучения для решения задач в экономике и финансах	Сформированное умение применять методы нейросетевого моделирования и глубокого обучения для решения задач в экономике и финансах
	Владеть: навыками нейросетевого моделирования и глубокого обучения для решения задач в экономике и финансах в современных средах	Фрагментарное владение навыками нейросетевого моделирования и глубокого обучения для решения задач в экономике и финансах в современных средах	В целом успешное, но не систематическое применение навыков нейросетевого моделирования и глубокого обучения для решения задач в экономике и финансах в современных средах	В целом успешное, но содержащее отдельные пробелы применение навыков нейросетевого моделирования и глубокого обучения для решения задач в экономике и финансах в современных средах	Успешное и систематическое применение навыков нейросетевого моделирования и глубокого обучения для решения задач в экономике и финансах в современных средах

Показатели сформированности компетенции:

Критериями оценивания являются баллы, которые выставляются преподавателем за виды деятельности (оценочные средства) по итогам изучения модулей (разделов дисциплины).

Шкалы оценивания:

от 45 до 59 баллов – «удовлетворительно»;

от 60 до 79 баллов – «хорошо»;

от 80 баллов – «отлично».

4.2. Контрольные задания или иные материалы, необходимые для оценивания результатов обучения по дисциплине. Методические материалы, определяющие процедуры оценивания результатов обучения по дисциплине

Код и наименование индикатора достижения компетенции	Результаты обучения	Оценочные средства
<i>ПК 1.1 Применяет подходы к проведению бизнес-анализа, работе с информацией и требованиями, оценке эффективности работы по бизнес-анализу и умеет обосновать их выбор</i>	<i>Знать: основные принципы использования машинного обучения для задач кредитного риск-менеджмента (PD, EAD, LGD) основные принципы использования моделей выживаемости для зада коллекторского скоринга</i>	Проверка ответов на вопросы, тестирование
	<i>Уметь: применять инструменты машинного обучения для задач кредитного риск-менеджмента (PD, EAD, LGD), для задач кредитного скоринга, для задач прогноза оттока клиентов, для задач выявления мошеннических транзакций</i>	Проверка кейсов 1-5
	<i>Владеть: Методами применения алгоритмов машинного обучения в риск-менеджменте с использованием современных информационных средств и сред</i>	Проверка кейсов 1-5 Индивидуальный проект
<i>ПК 1.2 Определяет связи и зависимости между элементами информации бизнес-анализа</i>	<i>Знать: основные методы нейросетевого моделирования и глубокого обучения</i>	Проверка ответов на вопросы, тестирование доклад
	<i>Уметь: Применять методы нейросетевого моделирования и глубокого обучения для решения задач в экономикк и финансах</i>	Проверка кейсов 3-4
	<i>Владеть: навыками нейросетевого моделирования и глубокого обучения для решения задач в экономике и финансах в современных средах</i>	Проверка кейсов 3-4

Вопросы и задания для самостоятельной работы

Примерные вопросы для экзамена:

- 1 Методы классификации для PD моделирования
- 2 Рейтингование заемщиков на основе PD моделирования
- 3 Методы регрессии для LGD моделирования
- 4 Методы регрессии для EAD моделирования
- 5 Методы классификации для стресс-тестирования в банках
- 6 Критерий Хосмера-Лемешова для кредитного скоринга,
- 7 Задача коллекторского скоринга на основе моделей выживаемости.
- 8 Выявление предикторов риска разорения на основе модели Кокса
- 9 Задача прогноза оттока клиентов на основе модели выживаемости (модели ускоренной жизни AFT)
- 10 Нейросетевое моделирование в экономике
- 11 Задача адаптивной фильтрации; Методы безусловной оптимизации;
- 12 Однослойный персептрон.
- 13 SVM (машина опорных векторов). Ядро – RBF – радиальная базисная функция для SVM
- 14 Применение алгоритма SVM для решения задач определения мошеннических транзакций
- 15 Структуры глубоких нейронных сетей структуры глубоких нейронных сетей
- 16 Обратное распространение ошибки:

Кейс-задания

Кейс-задача 1. Раздел дисциплины «PD-модели»

Цель: Рассчитать вероятности дефолта (PD) заёмщиков на основе модели деревьев принятия решений CART. Провести рейтингование заёмщиков.

Провести классификацию заемщиков, для решения используя классификатор деревьев решения, предварительно проверив валидность исходных данных. Выбрать признак и значения порога, по которому происходит оптимальное по заданному критерию разбиение в алгоритме CART по вероятности дефолта (PD). Задать максимальное число объектов в вершине-листа дерева, для определения критерия останова алгоритма. Построить алгоритм на размеченных данных по кредитному скорингу. Провести классификацию заемщиков на основе расчета вероятности дефолта. Провести рейтингование заёмщиков

1 Содержание практики

1. Пакет caret

- Кросс-валидация
- Гридсерч
- Построение моделей
- Метрики

2. Модель дерева решений

3 Задание

В практическом задании вы будете использовать классический набор по кредитному скорингу. Вам необходимо

предсказать какие из заемщиков разоряться, а какие нет. То есть вычислить вероятность дефолта.

2.1 Создайте новый скрипт RMarkdown.

1. Загрузите файл с данными
2. Разделите выборки на train/test с соотношением 0.8
3. Зависимая (целевая) переменная: Target

2.2 Моделирование

Для модели кредитного скоринга получите прогноз и матрицу сопряженности. Также для модели через caret выпишите какие гиперпараметры оказались наилучшими.

1. Сформируйте параметры для кросс-валидации:
 - метод: cv
 - число фолдов: 6
2. Модель дерева решений
3. Провести рейтингование заёмщиков

4 Оформление отчета

Отчет необходимо оформить в RMarkdown. Отключите warning и message в куске кода (чанке) с

подключением пакетов. Скрывать код (echo) из отчета не нужно.

Кейс-задача 2. Определение длины горизонта и EAD в кредитном риск-менеджменте

Цель: Информация о дефолте используется для оценки ожидаемых потерь, под которые банк осуществляет резервирование. В этом случае возникает вопрос, а на каком периоде смотреть выходы в дефолт? Важно ли это? С точки зрения бизнеса, важно понимать период планирования, с точки зрения моделей следует выбрать такой период, в котором будет охвачено не менее 80%-90% всех возможных выходов в дефолт для всех открытых и не находящихся в дефолте договоров на текущий момент времени. На основе модели регрессии по случайному лесу рассчитать:

Задание

1. Рассчитать длину горизонта вероятного дефолта

1.1 рассмотреть несколько временных срезов (поколений/когорт) на исторической выборке;

1.2 для каждого из поколений учесть все наблюдения из выборки для моделирования, не находящиеся в дефолте;

1.3 Для таких наблюдений построить кумулятивный график выхода в дефолт. Временной шаг – месяц. График может быть построен как в штуках, так и в деньгах;

1.4 Определить горизонт, как временной интервал, на котором выходят в дефолт 80% всех тех клиентов, которые вообще окажутся дефолте за весь доступный исторический период.

2 Рассчитать – CCF – credit conversion factor – коэффициент кредитной конверсии:

2.1 определить EAD – сумма средств, который клиент должен банку в момент дефолта на основе модели регрессии (модель случайного леса),

2.2. рассчитать CCF – часть доступных на момент наблюдения средств, которая будет использована клиентом к моменту дефолта.

Кейс-задача 3. Расчет LGD на основе модели регрессии

Цель: Определить сумму LGD – loss given default, ту часть EAD, которую банку всё-таки не удастся вернуть на основе модели регрессии (LASSO)

Задание

Та часть EAD, которую банку всё-таки не удалось вернуть, называется LGD – loss given default. Поскольку временной интервал, на котором происходит т.н. «восстановление» (возвращение долга), может варьироваться по клиентно, возникает необходимость определить длину горизонта восстановления, на котором будет рассчитываться целевая переменная. На длину горизонта влияет в первую очередь доступность достаточного временного периода в данных для моделирования.

1. На основе линейной регрессии определить сумму LGD, выбрав длину горизонта в диапазоне 3 года после дефолта.
2. Вычислить денежный поток от клиента (выплаты, реализация залога и др.), полученный на горизонте восстановления после дефолта. При расчете денежного потока, который идет на восстановление, использовать дисконтирование – механизм учета текущей стоимости денежных средств, полученных на горизонте восстановления.
3. Рассчитать суммы возмещения, затрат и продажи долга при дисконтировании, умножив на «фактор дисконтирования» с учетом процентной ставки по договору.

Кейс-задание 4 (модуль 1). Построение модели кредитного скоринга на основе моделей бинарной регрессии.

1. Провести предварительный анализ исходных данных. Исключить аномальные наблюдения (если такие есть), заполнить пропуски (если они имеются). Провести корреляционный анализ независимых переменных, исключив переменные, значительно коррелирующие с другими переменными ($>0,9$).

2. Построить статистически значимую модель бинарной регрессии для кредитного скоринга, оценив параметры методом максимального правдоподобия, применяя метод пошагового исключения, в которой все переменные будут статистически значимы. Подобрать функцию распределения, описывающую вероятность положительной альтернативы (предотвращение дефолта) между нормальным распределением (пробит), логистическим (логит) и экстремальным (гомпит) на основе минимума информационных критериев.

3. Проверить качество отобранной модели, подтвердив его значениями коэффициентов R^2 МакФаддена, тестом отношения правдоподобия (LR-тестом), результатами теста Хосмера-Лемешоу и любым тестом на нормальность распределения остатков (например, Колмогорова-Смирнова или Бера-Жарка).

4. Рассчитать маргинальные эффекты и провести интерпретацию коэффициентов модели.

5. Оформить отчет о выполнении задания с приведением условия задачи, результатов решения и выводов.

1 Содержание практики

1. Оформление отчетов в RMarkdown
2. Формат данных .rds
3. Логит, пробит и гомпит модели
4. Селекция моделей (AIC, BIC)
5. Коэффициенты, графики и тесты
6. Маргинальные эффекты
7. Прогнозирование и матрицы сопряженности

2 Задание

Вам предстоит определить какие показатели влияют на удовлетворённость жизнью

2.1 Создайте новый скрипт RMarkdown

1. Загрузите файл Credit.rds
 2. Разделите выборки на train/test с соотношением 0.8
 3. Зависимая (целевая) переменная: Target
- ### 2.2 Моделирование
- Вам необходимо построить логит, пробит и гомпит модели. Для каждого типа вам необходимо повторить следующие шаги:

1. Постройте модель на всех переменных
2. Проверьте какие переменные оказались не значимыми
3. Перестройте модель без эти переменных
4. Повторяйте действия пока все переменные не будут значимыми

2.3 Селекция и тесты

1. С помощью информационных критериев (AIC, BIC) выберите лучшую модель
2. Тесты и коэффициенты: коэффициент детерминации Макфаддена, Likelihood-ratio, Колмогоров- Смирнов, Хосмер-Лемешоу, графики распределения ошибок
3. Проведите анализ тестов

2.4 Прогнозирование и маргинальные эффекты

1. Получите прогноз
2. Сформируйте матрицу сопряженности
3. Посчитайте маргинальные эффекты

Примечание: В пакете `mfX` нет маргинальных эффектов для гомпит-модели, если лучшей моделью окажется гомпит-модель, получите маргинальные эффекты для логит или пробит модели. В зависимости от того, какая из моделей будет лучше по AIC и BIC

3 Оформление отчета

Отчет необходимо оформить в RMarkdown. Отключите `warning` и `message` в куске кода (чанке) с подключением пакетов. Скрывать код (`echo`) из отчета не нужно.

Кейс-задание 5 (модуль 2). Построение модели коллекторского скоринга на основе моделей выживаемости (Кокса).

2. Загрузите данные Landing Clab по коллекторскому скорингу.
3. Постройте графики функций выживаемости (времени до дефолта), оцененные методом Каплана–Майера.
Проанализируйте полученные результаты, сделав предварительный вывод о различии функций выживаемости для разных категорий заемщиков
4. Проведите логранговый тест Мантеля–Хензеля и тест Гехана–Вилкоксона на значимое различие в оценках выживаемости по подгруппам, сформированным по атрибутивному признаку (назначение кредита).
5. Постройте модель пропорциональных рисков Кокса по риску разорения, для оценки коэффициентов модели использовать метод частичного правдоподобия согласно методикам Эфрона или Брэслоу.
6. Оцените качество моделей на основе показателя R^2_{mer} (мера объясненной случайности,) и R^2_{mev} (мера объясненного отклонения), выберите лучшую методику оценки исходя из минимума информационных критериев Акайке и Шварца. Проведите интерпретацию коэффициентов отобранной модели с точки зрения предикторов риска разорения.
7. Оформите отчёт в соответствии с шаблоном.

Кейс-задание 6 (модуль 3). Построение модели прогнозирования мошеннических транзакций на основе машины опорных векторов. Виртуальный кластер.

1. Модель SVM (Машина опорных векторов)
2. Параллелизация вычислений для их ускорения
 - создание виртуальных вычислительных кластеров
 - построение моделей с параллелизацией вычислений
 - сравнение времени, которое требуется на построение моделей с/без параллелизации вычислений.

2 Задание

В практическом задании следует продолжить работать с данными транзакций кредитных карт. Вычисление на большом количестве данных может занять значительное время, некоторые алгоритмы (например, XGBoost) автоматически используют весь ЦП, но другим требуется настроить параллелизацию вычислений (например, SVM)

2.1 Создайте новый скрипт RMarkdown.

1. Загрузите файл `credit.rds` или `credit_light.rds` если у вас не получится работать со стандартным набором
2. Разделите выборки на обучающую (`train`) и тестовую (`test`) с соотношением 0.75
3. Зависимая (целевая) переменная: `Target`
4. Выполните задания по моделированию с помощью метода машины опорных векторов SVM
5. Скомпилируйте отчет в `html` или `docx`

2.2 Моделирование

1. Построить модель SVM на данных о транзакциях без использования пакета `caret`, а с использованием пакета `kernlab`. В качестве функции ядра выбрать радиальную базисную функцию `kernel = "rbfdot"`. Дополнительно можно построить модель с линейным ядром, то есть выбрать `kernel = "vanilladot"`.
2. Построить модель SVM с `caret`, но без параллелизации: для формирования сетки поиска (`GridSearch`) выбрать гиперпараметры `svm_grid = expand.grid(C = c(5, 6), sigma = c(0.03, 0.035))`
3. Сравнить результаты построенных моделей и сделать выводы
4. Проверить сколько времени уходит на обычное построение модели без проведения параллелизации
5. Создать виртуальный вычислительный кластер
6. Проверить сколько времени уходит на построение модели с параллелизацией
7. Сравните результаты

Темы докладов:

1. применение искусственного интеллекта в рамках программы «Умный город».
2. применение искусственного интеллекта в финансах
3. применение искусственного интеллекта в ЖКХ
4. применение искусственного интеллекта в управлении проектами
5. применение искусственного интеллекта в макроэкономическом прогнозировании
6. применение искусственного интеллекта в поведенческой экономике

Требование к докладу

1. Доклад должен отражать только актуальные тенденции интеллектуального анализа в

экономике

2. При подготовке к докладу должно быть использовано не менее 5 источников информации
3. Доклад должен занимать по объему от 6-8 минут, и 2 минуты отводится на дискуссию и вопросы
4. Доклад должен сопровождаться презентацией 7-10 слайдов

Критерии оценки докладов:

- оценка «зачтено» выставляется магистранту, при выполнении всех требований к докладу, а также умение аргументировано отвечать на задаваемые по теме доклада вопросы.
- оценка «не зачтено» выставляется магистранту, при невыполнении всех требований к докладу, а также неумение аргументировано отвечать на задаваемые по теме доклада вопросы.

Экзаменационные билеты

Экзамен является оценочным средством для всех этапов освоения компетенций.
В экзаменационном билете – 2 вопроса .

Образец экзаменационного билета:

Башкирский государственный университет	Направление подготовки 38.04.05 «Бизнес-информатика»
Институт экономики, финансов и бизнеса	Программа «Информационная бизнес-аналитика»
Кафедра математических методов в экономике	Дисциплина «Анализ неструктурированной информации»

Билет к экзамену № 1

1. Рейтингование заемщиков на основе PD моделирования
2. Применение алгоритма SVM для решения задач определения мошеннических транзакций

Заведующий кафедрой

Р.Х. Бахитова

5. Учебно-методическое и информационное обеспечение дисциплины

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

Основная литература

1. Анализ данных : учебник для академического бакалавриата / ГУ - Высшая школа экономики; под ред. В. С. Мхитаряна .— Москва : Юрайт, 2016 .— 490 (13 экз.)
2. Новиков, Федор Александрович. Символический искусственный интеллект: математические основы представления знаний : учеб. пособие для академического бакалавриата / Ф. А. Новиков .— Москва : Юрайт, 2019 .— 278 с. :
3. Искусственный интеллект и интеллектуальные системы управления / И. М. Макаров [и др.] ; РАН, Отделение информационных технологий и вычислительных систем .— М. : Наука, 2006 .— 333 с.
4. Кондрашов, Юрий Николаевич. Анализ данных и машинное обучение на платформе MS SQL Server : учебное пособие / Ю. Н. Кондрашов ; Финансовый университет при Правительстве РФ ; Департамент анализа данных, принятия решений и финансовых технологий .— Москва : РУСАЙНС, 2020 .— 303 с.
5. Криптоэкономика: мошенников выявит искусственный интеллект [[Текст]] // Эксперт. — 2018 .— № 43 .— С. 7

Дополнительная литература

1. Чернов, Владимир Анатольевич . Интегрированное информационно-аналитическое обеспечение и искусственный интеллект в системе управления [[Текст]] = The integrated information and analytical providing and artificial intelligence in a control system / Чернов Владимир Анатольевич // Аудит. — 2017 .— № 1 .— С. 17-23
2. Бутенко, Екатерина Дмитриевна . Искусственный интеллект в банках сегодня: опыт и перспективы [[Текст]] / Е. Д. Бутенко // Финансы и кредит. — 2018 .— Т. 24, вып. 1 .— С. 143-153
3. Хорошилов, Евгений Евгеньевич. Новые технологии в финансовом секторе Канады: искусственный интеллект и "большие данные" [[Текст]] = New Technologies in Canadian Financial Sector: Artificial Intelligence and Big Data / Е. Е. Хорошилов // США. Канада. Экономика - политика - культура. — 2018 .— № 10 .— С. 50-65.
4. Винстон, Уэйн. Бизнес-моделирование и анализ данных. Решение актуальных задач с помощью Microsoft EXCEL : пер. с англ. яз. / У. Винстон ; перевод Ю. Бочиной .— 5-е изд. — Санкт-Петербург : Питер, 2018 .— 864 с. (8 экз.)

5.2. Перечень ресурсов информационно-телекоммуникационной сети «Интернет» и программного обеспечения, необходимых для освоения дисциплины

Пользователям библиотеки БашГУ предоставляется возможность использования следующих электронных информационных ресурсов:

№	Наименование Интернет-ресурса	Ссылка (URL) на Интернет ресурс
1.	Федеральная служба государственной статистики	www.gks.ru
2.	Министерство финансов РФ	www.minfin.ru
3.	Международный валютный фонд	www.imf.org
4.	Центр макроэкономического анализа и краткосрочного прогнозирования	www.forecast.ru
5.	Территориальный орган Федеральной службы государственной статистики по РБ	www.bashstat.ru
6.	Информационно-издательский центр «Статистика России»	www.infostat.ru
7.	Информационно-аналитический сайт в области информационных технологий	citforum.ru
8.	Издание о высоких технологиях	cnews.ru
9.	Библиотека Г. Верникова – все о менеджменте и ИТ - подборка аналитических материалов по вопросам экономики, менеджмента и информационных технологий.	vernikov.ru
10.	Официальный портал ИТ-директоров (Реестр ИТ-поставщиков)	globalcio.ru
11.	Журнал СІО – руководитель информационной службы	cio-world.ru
12.	Единый архив экономических и социологических данных ВШЭ	http://sophist.hse.ru/

1. База данных периодических изданий на платформе EastView: «Вестники Московского университета», «Издания по общественным и гуманитарным наукам» - <https://dlib.eastview.com/>

2. Информационная система «Единое окно доступа к образовательным ресурсам» - <http://window.edu.ru>

3. Научная электронная библиотека eLibrary.ru - <http://elibrary.ru/defaultx.asp>

4. Справочно-правовая система Консультант Плюс - <http://www.consultant.ru/>

5. Электронная библиотечная система «Университетская библиотека онлайн» – <https://biblioclub.ru/>

6. Электронная библиотечная система «ЭБ БашГУ» – <https://elib.bashedu.ru/>

7. Электронная библиотечная система издательства «Лань» – <https://e.lanbook.com/>

8. Электронный каталог Библиотеки БашГУ – <http://www.bashlib.ru/catalogi>.

9. Архивы научных журналов на платформе НЭИКОН (Cambridge University Press, SAGE Publications, Oxford University Press) - <https://archive.neicon.ru/xmlui/>
10. Издательство «Annual Reviews» - <https://www.annualreviews.org/>
11. Издательство «Taylor&Francis» - <https://www.tandfonline.com/>
12. Windows 8 Russian. Windows Professional 8 Russian Upgrade Договор №104 от 17.06.2013 г. Лицензии бессрочные.
13. Microsoft Office Standard 2013 Russian. Договор №114 от 12.11.2014 г. Лицензии бессрочные.
14. Windows 8 Russian. Windows Professional 8 Russian Upgrade. Договор № 104 от 17.06.2013 г. Лицензии бессрочные.

6. Материально-техническая база, необходимая для осуществления образовательного процесса по дисциплине

Наименование специализированных аудиторий, кабинетов, лабораторий	Вид занятий	Наименование оборудования, программного обеспечения
<p><i>учебная аудитория для проведения занятий лекционного типа:</i> лаборатория социально-экономического моделирования № 107 (помещение, ул.Карла Маркса, д.3, корп.4), лаборатория анализа данных № 108 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 110 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 111 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 114 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 122 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 204 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 207 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 208 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 209 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 210 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 212 (гуманитарный корпус), аудитория № 213 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 218 (гуманитарный корпус), аудитория № 220 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 221 (гуманитарный корпус), аудитория № 222 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 301 (гуманитарный корпус), аудитория № 305 (гуманитарный корпус), аудитория № 307 (гуманитарный корпус), аудитория № 308 (гуманитарный корпус), аудитория № 309 (гуманитарный корпус), аудитория № 110 (гуманитарный корпус), лаборатория исследования процессов в экономике и управлении № 311а (гуманитарный корпус), лаборатория информационных технологий в экономике и управлении № 311в (гуманитарный корпус).</p>	<p>Лекции</p>	<p>лаборатория социально-экономического моделирования № 107: учебная мебель, доска, проекционный экран с светодиодом lumien master control, проектор casio, персональный компьютер пэвм кламас в комплекте – 18 шт.</p> <p>лаборатория анализа данных № 108: учебная мебель, доска, персональный компьютер пэвм кламас в комплекте – 17 шт.</p> <p>аудитория № 110: учебная мебель, доска, телевизор led.</p> <p>аудитория № 111: учебная мебель, доска, телевизор led.</p> <p>аудитория № 114: учебная мебель, доска.</p> <p>аудитория № 115: учебная мебель, колонки (2 шт.), динамики, dvd плеер toshiba, магнитола sony (4 шт.)</p> <p>аудитория №118: учебная мебель, проектор benq, колонки (2 шт.), музыкальный центр lg, флипчарт магнитно-маркерный на треноге</p> <p>аудитория № 122: учебная мебель, доска.</p> <p>аудитория № 204: учебная мебель, доска, проекционный экран с светодиодом lumien master control, проектор casio.</p> <p>аудитория № 207: учебная мебель, доска, телевизор led tcl.</p> <p>аудитория № 208: учебная мебель, доска, телевизор led tcl.</p> <p>аудитория № 209: учебная мебель, доска.</p>
<p><i>учебная аудитория для проведения занятий семинарского типа:</i> лаборатория социально-экономического моделирования № 107 (помещение, ул.Карла Маркса, д.3, корп.4), лаборатория анализа данных № 108 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 110 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 111 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 114 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 122 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 204 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 207</p>	<p>Практические/семинарские занятия</p>	<p>аудитория № 210: учебная мебель, доска.</p> <p>аудитория № 212: учебная мебель, доска, проектор infocus.</p> <p>аудитория № 213: учебная мебель, доска, проекционный экран с светодиодом lumien master control, проектор casio.</p> <p>аудитория № 218: учебная мебель, доска, мультимедиа-проектор infocus.</p> <p>аудитория № 220:</p>

<p>(помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 208 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 209 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 210 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 212 (гуманитарный корпус), аудитория № 213 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 218 (гуманитарный корпус), аудитория № 220 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 221 (гуманитарный корпус), аудитория № 222 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 301 (гуманитарный корпус), аудитория № 305 (гуманитарный корпус), аудитория № 307 (гуманитарный корпус), аудитория № 308 (гуманитарный корпус), аудитория № 309 (гуманитарный корпус), аудитория № 110 (гуманитарный корпус), лаборатория исследования процессов в экономике и управлении № 311а (гуманитарный корпус), лаборатория информационных технологий в экономике и управлении № 311в (гуманитарный корпус).</p>		<p>учебная мебель, доска. аудитория № 221 учебная мебель, доска. аудитория № 222 учебная мебель, доска. аудитория № 301 учебная мебель, экран на штативе, проектор aser. аудитория № 302 учебная мебель, персональный компьютер в комплекте hp, моноблок, персональный компьютер в комплекте моноблок iru. аудитория № 305 учебная мебель, доска, проектор infocus. аудитория № 307 учебная мебель, доска. аудитория № 308 учебная мебель, доска. аудитория № 309 учебная мебель, доска. лаборатория исследования процессов в экономике и управлении № 311а учебная мебель, доска, персональный компьютер lenovo thinkcentre – 16 шт. лаборатория информационных технологий в экономике и управлении № 311в учебная мебель, доска, персональный компьютер в комплекте № 1 iru corp 510 – 14 шт. аудитория № 312 учебная мебель, доска.</p>
<p>учебная аудитория для проведения групповых и индивидуальных консультаций: лаборатория социально-экономического моделирования № 107 (помещение, ул.Карла Маркса, д.3, корп.4), лаборатория анализа данных № 108 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 110 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 111 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 114 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 122 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 204 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 207 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 208 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 209 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 210 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 212 (гуманитарный корпус), аудитория № 213 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 218 (гуманитарный корпус), аудитория № 220 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 221 (гуманитарный корпус), аудитория № 222 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 301 (гуманитарный корпус), аудитория № 305 (гуманитарный корпус), аудитория № 307 (гуманитарный корпус), аудитория № 308 (гуманитарный корпус), аудитория № 309 (гуманитарный корпус), аудитория № 110 (гуманитарный корпус), лаборатория исследования процессов в экономике и управлении № 311а</p>	<p>Групповые и индивидуальные консультации</p>	<p>1. Windows 8 Russian. Windows Professional 8 Russian Upgrade. Договор № 104 от 17.06.2013 г. Лицензии – бессрочные. Microsoft Office Standard 2013 Russian. Договор № 114 от 12.11.2014 г. Лицензии – бессрочные.</p>

(гуманитарный корпус), лаборатория информационных технологий в экономике и управлении № 311в (гуманитарный корпус).		
<p>учебная аудитория для текущего контроля и промежуточной аттестации: лаборатория социально-экономического моделирования № 107 (помещение, ул.Карла Маркса, д.3, корп.4), лаборатория анализа данных № 108 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 110 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 111 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 114 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 122 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 204 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 207 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 208 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 209 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 210 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 212 (гуманитарный корпус), аудитория № 213 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 218 (гуманитарный корпус), аудитория № 220 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 221 (гуманитарный корпус), аудитория № 222 (помещение, ул.Карла Маркса, д.3, корп.4), аудитория № 301 (гуманитарный корпус), аудитория № 305 (гуманитарный корпус), аудитория № 307 (гуманитарный корпус), аудитория № 308 (гуманитарный корпус), аудитория № 309 (гуманитарный корпус), аудитория № 110 (гуманитарный корпус), лаборатория исследования процессов в экономике и управлении № 311а (гуманитарный корпус), лаборатория информационных технологий в экономике и управлении № 311в (гуманитарный корпус).</p>	Текущий контроль и промежуточная аттестация	
<p>помещения для самостоятельной работы: аудитория № 302 читальный зал (гуманитарный корпус).</p>	Самостоятельная работа	