

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Волков В.В.
Должность: Ректор
Дата подписания: 15.07.2024 14:54:55
Уникальный программный ключ:
ed68fd4b85b778e0f0b1bfea5dbc56cf4148f1229917e799a70e51517ff6d591

Автономная некоммерческая образовательная организация высшего образования
«Европейский университет в Санкт-Петербурге»

УТВЕРЖДАЮ

Ректор  / В.В. Волков

« 28 » февраля 2024 г.

Протокол Ученого Совета

№ 2 от 28 февраля 2024 г.



Рабочая программа дисциплины
«Машинное обучение»

дополнительная профессиональная программа
«Прикладной анализ данных»

вид программы
программа профессиональной переподготовки

язык обучения – русский
форма обучения – очная

Санкт-Петербург

Авторы:

Тушканова О.Н., кандидат технических наук, доцент факультета социологии АНООВО «ЕУСПб»

Рабочая программа дисциплины «Машинное обучение», входящая в состав дополнительной профессиональной программы профессиональной переподготовки «Прикладной анализ данных» утверждена на заседании Ученого совета университета.

Содержание

1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ. ОБЪЕМ ДИСЦИПЛИНЫ.....	4
2. ТРЕБОВАНИЯ К УРОВНЮ ОСВОЕНИЯ СОДЕРЖАНИЯ ДИСЦИПЛИНЫ.....	4
3. СОДЕРЖАНИЕ И СТРУКТУРА ДИСЦИПЛИНЫ.....	5
4. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ САМОСТОЯТЕЛЬНОЙ РАБОТЫ	6
5. ОЦЕНОЧНЫЕ СРЕДСТВА.....	7
6. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ.....	8
7. ПРОГРАММНОЕ И МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ПРОГРАММЫ.....	10

1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ. ОБЪЕМ ДИСЦИПЛИНЫ

Цель обучения: освоения дисциплины «Машинное обучение» — знакомление с типами задач анализа данных и методов анализ данных.

Задачи обучения:

- применение алгоритма машинного обучения в соответствии с поставленной задачей анализа данных;
- реализация программ с использованием современных методов и библиотек машинного обучения.

Изучение данной дисциплины способствует формированию профессиональных навыков по разработке теоретических моделей исследуемых процессов, явлений и объектов, относящихся к сфере профессиональной деятельности, их исследованию, оценке и интерпретации полученных результатов.

Отличительной особенностью реализуемого подхода к преподаванию дисциплины является разнообразных практических иллюстраций основных теоретических положений.

Общая трудоемкость дисциплины составляет 62 часа.

2. ТРЕБОВАНИЯ К УРОВНЮ ОСВОЕНИЯ СОДЕРЖАНИЯ ДИСЦИПЛИНЫ

В результате освоения дисциплины слушатель должен приобрести следующие знания и умения, необходимые для качественного изменения и (или) получения новых профессиональных компетенций:

слушатель должен знать:

- особенности применения различных теоретико-методологических концепций с использованием технологий прикладного анализа данных;
- библиотеки машинного обучения.

слушатель должен уметь:

- уметь применять алгоритмы машинного обучения;
- правильно оформлять и представлять результаты работы.

слушатель должен владеть:

- навыками реализация программ и применение современных методов и библиотек машинного обучения;
- способностью к самостоятельному обучению новым методам исследования, к изменению научного и научно-производственного профиля своей профессиональной деятельности.

В результате изучения дисциплины «Машинное обучение» слушатель приобретает следующие профессиональные компетенции (Таблица 1):

Планируемые результаты обучения по дисциплине

Таблица 1

Код и название компетенции	Содержание компетенции	Планируемые результаты обучения по дисциплине, характеризующие этапам формирования компетенций
ОПК-1	способен использовать современные информационные технологии и программные средства, при решении задач профессиональной деятельности	Знать: З (ОПК-1) – современные информационные технологии и программные средства, в том числе отечественного производства при решении задач профессиональной деятельности.
		Уметь: У (ОПК-1) – выбирать современные информационные технологии и программные средства, в том числе отечественного производства при решении задач профессиональной деятельности

Код и название компетенции	Содержание компетенции	Планируемые результаты обучения по дисциплине, характеризующие этапам формирования компетенций
		Владеть: В (ОПК-1) - навыками применения современных информационных технологий и программных средств, при решении задач профессиональной деятельности
ОПК-2	способен принимать участие в реализации профессиональных коммуникаций с заинтересованными участниками проектной деятельности и в рамках проектных групп	Уметь: У (ОПК-2) – осуществлять взаимодействие с заказчиком в процессе реализации проекта; принимать участие в командообразовании и развитии персонала
		Владеть: В (ОПК-2) – навыками проведения презентаций, переговоров, публичных выступлений
ПК-6	способен применить технологии машинного обучения к реальным общественным задачам	Знать: З (ПК-6) – основы технологий машинного обучения
		Уметь: У (ПК-6) – анализировать текущие проблемы социальных и общественных наук
		Владеть: В (ПК-6) – навыками применения машинного обучения к реальным общественным задачам

3. СОДЕРЖАНИЕ И СТРУКТУРА ДИСЦИПЛИНЫ

Содержание дисциплины соотносится с планируемыми результатами обучения по дисциплине через задачи, формируемые компетенции и их компоненты (знания, умения, навыки – далее ЗУВ) в соответствии с таблицей 2.

Таблица 2

Содержание дисциплины

№ п/п	Наименование тем (разделов)	Содержание тем (разделов)	Коды компетенций	Коды ЗУВ (в соответствии с табл. 1)
1	SVM	SVM. Прямая и двойственная задача. Решение двойственной задачи. Линейно неразделимый случай. Ядерный трюк.	ОПК-1 ПК-6	З (ОПК-1) У (ОПК-1) В (ОПК-1) З (ПК-6) У (ПК-6) В (ПК-6)
2	Байес. Байесовский классификатор	Оценка распределения признаков. EM-алгоритм. Регрессия. Линейная регрессия, гребневая регрессия, LASSO, LARS. Борьба с выбросами.	ОПК-1 ПК-6	З (ОПК-1) У (ОПК-1) В (ОПК-1) З (ПК-6) У (ПК-6)
3	Локальный поиск	Байесовская оптимизация. Hill Climb, отжиг, генетический алгоритм.	ОПК-1 ПК-6	З (ОПК-1) У (ОПК-1) В (ОПК-1) З (ПК-6) У (ПК-6)
4	Обучение с подкреплением	Постановка задачи. Функция ценности и функция качества. Q-обучение. Глубокое обучение с подкреплением, DQN, Actor-Critic.	ОПК-1 ОПК-2 ПК-6	З (ОПК-1) У (ОПК-1) В (ОПК-1) У (ОПК-2) В (ОПК-2) З (ПК-6) У (ПК-6)

Структура дисциплины

Таблица 3.

№ п/п	Наименование и содержание тем	Основные понятия (категории) и проблемы, рассматриваемые в теме	Объем дисциплины, час.				Форма текущего контроля успеваемости, промежуточной аттестации
			Всего	Аудиторная работа по видам учебных занятий		СР ¹	
				Л	СЗ ²		
1.	SVM	SVM. Прямая и двойственная задача. Решение двойственной задачи. Линейно неразделимый случай. Ядерный трюк.	15	4	3	8	Практические задания
2.	Байес. Байесовский классификатор	Оценка распределения признаков. EM-алгоритм. Регрессия. Линейная регрессия, гребневая регрессия, LASSO, LARS. Борьба с выбросами.	15	3	4	8	
3.	Локальный поиск	Байесовская оптимизация. Hill Climb, отжиг, генетический алгоритм.	15	4	3	8	
4.	Обучение с подкреплением	Постановка задачи. Функция ценности и функция качества. Q-обучение. Глубокое обучение с подкреплением, DQN, Actor-Critic.	15	3	4	8	
5.	Промежуточная аттестация	Опрос по вопросам (устно)	2	-	-	2	
Всего:			62	14	14	34	

4. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ САМОСТОЯТЕЛЬНОЙ РАБОТЫ

Общие положения.

Знания и навыки, полученные в результате лекций и семинарских занятий, закрепляются и развиваются в результате повторения материала, усвоенного в аудитории, путем чтения исследовательской литературы (из списков основной, дополнительной), статей по проблематике занятия и их анализа.

Самостоятельная работа обучающегося представляет самостоятельное изучение дополнительных материалов, Интернет-ресурсов и пр. Подготовка к семинарским занятиям, выполнение практических заданий, создание докладов, проектов и презентаций также является важной формой работы обучающихся. Самостоятельная работа может вестись как индивидуально, так и при содействии преподавателя. Вопросы и замечания, возникшие в ходе самостоятельного внеаудиторного чтения рекомендованной литературы, обсуждаются с преподавателем и другими обучающимися. Выносятся на обсуждение, как правило, актуальные проблемы и предлагается их рассмотреть с точки зрения того или иного теоретического подхода.

На занятиях материал излагается в проблемной форме. Основной упор в преподавании делается на изучение теоретических понятий и возможности их применения на конкретных примерах, в том числе в устных выступлениях обучающихся.

¹ Самостоятельная работа, включает в себя часы на промежуточный контроль

² Могут включать в себя: лабораторные работы, круглые столы, мастер-классы, мастерские, деловые игры, ролевые игры, тренинги, семинары по обмену опытом, выездные занятия, консультации

Перечень основных вопросов по изучаемым темам для самостоятельной работы обучающихся.

- Байесовский классификатор. Типы оценки распределений признаков (Gaussian, Bernoulli, Multinomial). EM алгоритм.
- Нейронные сети. Перцептрон Розенблатта. Функции активации. Обратное распространение градиента. Softmax.
- Стохастическая оптимизация. Hill Climb. Отжиг. Генетический алгоритм.

Источники для самостоятельной подготовки:

- Кук, Д. Машинное обучение с использованием библиотеки H2O / Д. Кук; пер. с англ. А.Б. Огурцова. - Москва: ДМК Пресс, 2018. - 250 с. - ISBN 978-5-97060-508-0. - Текст: электронный. - URL: <https://znanium.com/catalog/product/1028135>
- Обработка изображений с помощью OpenCV / Глория Буэно Гарсия [и др.]; пер. с англ. А.А. Слинкина. - Москва: ДМК Пресс, 2016. - 210 с. - ISBN 978-5-970(0)-387-1. - Текст: электронный. - URL: <https://znanium.com/catalog/product/1028080>
- Груздев, А.В. Прогнозное моделирование в IBM SPSS Statistics и R: Метод деревьев решений / А.В. Груздев. - Москва: ДМК Пресс, 2016. - 278 с. - ISBN 978-5-97060-456-4. - Текст: электронный. - URL: <https://znanium.com/catalog/product/1028064>
- Горелов, В.И. Анализ статистических данных: практикум: [16+] / В.И. Горелов, Т.Н. Ледацева; Российская международная академия туризма. – Москва: Университетская книга, 2015. – 120 с.: ил. – Режим доступа: по подписке. – URL: <http://biblioclub.ru/index.php?page=book&id=574944>

5. ОЦЕНОЧНЫЕ СРЕДСТВА

Проведение текущего контроля в рамках реализации данной дисциплины проходит в соответствии с Таблицей 3 данной рабочей программы дисциплины по основным понятиям (категориям) и проблемам, рассматриваемым в предложенных темах. Фиксация результатов текущего контроля в рамках реализации данной дисциплины не предусмотрена.

Типовые задания к текущей аттестации (практические задания).

Практическое задание 1:

- Реализуйте алгоритм линейной регрессии, и полиномиальной регрессии (для датасета `poisysine` – степеней от 2 до 5, для датасета `hydrodynamics` – степени 2) без регуляризации.
- Реализуйте алгоритм гребневой регрессии и найдите оптимальный параметр регуляризации для случаев из задачи 1.
- Найдите максимум функции с помощью алгоритма кросс-энтропийного поиска, изображая распределение на каждом шаге.
- Найдите лучший путь в задаче коммивояжера с помощью алгоритма отжига.

Критерии оценивания

Формы текущего контроля успеваемости	Критерии оценивания
Практическое задание	к содержанию, структуре, логике, аргументации, оформлению с возможным небольшим количеством погрешностей (например, плохо выдержанная структура текста, недостаточная аргументация отдельных тезисов) – зачтено слушатель выполняет задание частично или с существенными недочетами (некорректно сформулирован исследовательский вопрос, не определены основные агенты, некорректно выбраны методы исследования, требования к содержанию, структуре, логике, аргументации, оформлению не выполнены) – не зачтено, полное и правильное выполнение задания в соответствии с требованиями

Форма промежуточной аттестации – зачет, выставляемый на основе устного ответа на вопросы.

При аттестации используются система «зачтено» и «не зачтено» в соответствии с критериями оценивания.

В результате промежуточного контроля знаний обучающиеся получают аттестацию по дисциплине.

Показатели, критерии и оценивание компетенций по уровням их формирования в процессе промежуточной аттестации

Таблица 4

Форма промежуточной аттестации/вид промежуточной аттестации	Коды компетенций	Коды ЗУВ (в соответствии с Таблицей 1)	Критерии оценивания	Оценка
зачет / устный ответ на вопросы	ОПК-1 ОПК-2 ПК-6	З (ОПК-1) У (ОПК-1) В (ОПК-1) У (ОПК-2) В (ОПК-2) З (ПК-6) У (ПК-6)	Слушатель верно отвечает на вопрос, указанный в билете, при условии, что ответ на вопрос характеризуется отсутствием серьезных, значимых неточностей, при следующих характеристиках ответа: <ul style="list-style-type: none"> – твердое знание материала курса, – последовательное изложение материала, – знание теоретических положений без обоснованной их аргументации, – соблюдение норм устной и письменной литературной речи. 	зачтено
			Слушатель представляет ответ на вопрос билета, свидетельствующий о некомпетентности слушателя, при следующих параметрах ответа: <ul style="list-style-type: none"> – незнание значительной части программного материала, – наличие существенных ошибок в определениях, формулировках, понимании теоретических положений; – бессистемность при ответе на поставленный вопрос, – отсутствие в ответе логически корректного анализа, аргументации, классификации, – наличие нарушений норм устной и письменной литературной речи. 	не зачтено

Результаты промежуточного контроля по дисциплине, выраженные в бинарной системе «зачтено», показывают уровень сформированности у обучающегося компетенций.

Результаты промежуточного контроля по дисциплине, выраженные в бинарной системе «не зачтено», показывают не сформированность у обучающегося компетенций по дисциплине.

Типовые вопросы к промежуточной аттестации.

- Размерность Вапника-Червоненкиса. Размерность Вапника-Червоненкиса для перцептрона.
- Логистическая регрессия. Градиентный спуск.
- Пороговые условия. Эффективность по Парето. Precision-Recall и ROC кривые. AUC.
- Ансамблевые методы регрессии. RANSAC. Theil-Sen. Huber.
- Перцептрон. Перцептрон с карманом.

- Метод опорных векторов. Постановка задачи. Формулировка и решение двойственной задачи. Типы опорных векторов. Ядра.
- Гипотезы и дихотомии. Функция роста. Точка поломки. Доказательство полиномиальности функции роста в присутствии точки поломки.
- Деревья решений. Информационный выигрыш, критерий Джини. Регуляризация деревьев. Небрежные решающие деревья.
- Байесовский классификатор. Типы оценки распределений признаков (Gaussian, Bernoulli, Multinomial). EM алгоритм.
- Нейронные сети. Перцептрон Розенблатта. Функции активации. Обратное распространение градиента. Softmax.
- Стохастическая оптимизация. Hill Climb. Отжиг. Генетический алгоритм.
- Метрические классификаторы. kNN. WkNN. Отбор эталонов. DROP5. Kdtree

6. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Основная литература:

- Кук, Д. Машинное обучение с использованием библиотеки H2O / Д. Кук; пер. с англ. А.Б. Огурцова. - Москва: ДМК Пресс, 2018. - 250 с. - ISBN 978-5-97060-508-0. - Текст: электронный. - URL: <https://znanium.com/catalog/product/1028135>
- Обработка изображений с помощью OpenCV / Глория Буэно Гарсия [и др.] ; пер. с англ. А.А. Слинкина. - Москва: ДМК Пресс, 2016. - 210 с. - ISBN 978-5-970(0)-387-1. - Текст: электронный. - URL: <https://znanium.com/catalog/product/1028080>

Дополнительная литература:

- Груздев, А.В. Прогнозное моделирование в IBM SPSS Statistics и R: Метод деревьев решений / А.В. Груздев. - Москва: ДМК Пресс, 2016. - 278 с. - ISBN 978-5-97060-456-4. - Текст: электронный. - URL: <https://znanium.com/catalog/product/1028064>
- Горелов, В.И. Анализ статистических данных: практикум: [16+] / В.И. Горелов, Т.Н. Ледащева; Российская международная академия туризма. – Москва: Университетская книга, 2015. – 120 с.: ил. – Режим доступа: по подписке. – URL: <http://biblioclub.ru/index.php?page=book&id=574944>.

Перечень информационно-справочных систем и профессиональных баз данных информационно-телекоммуникационной сети «Интернет», необходимых для освоения программы:

Информационно-справочные системы:

- Гарант.Ру. Информационно-правовой портал: <http://www.garant.ru>
- Информационная система «Единое окно доступа к образовательным ресурсам»: <http://window.edu.ru/>
- Открытое образование. Ассоциация «Национальная платформа открытого образования»: <http://npoed.ru>
- Официальная Россия. Сервер органов государственной власти Российской Федерации: <http://www.gov.ru>
- Официальный интернет-портал правовой информации. Государственная система правовой информации: <http://pravo.gov.ru>
- Правовой сайт КонсультантПлюс: <http://www.consultant.ru/sys>
- Российское образование. Федеральный портал: <http://www.edu.ru>

Тематические системы:

- Google. Книги: <https://books.google.com>
- Internet Archive: <https://archive.org>
- Koob.ru. Электронная библиотека «Куб»: <http://www.koob.ru/philosophy/>

- Библиотека Ихтика [ihtik.lib.ru]: <http://ihtik.lib.ru/>
- Докусфера — Российская национальная библиотека: <http://leb.nlr.ru>
- ЕНИП — Электронная библиотека «Научное наследие России»: <http://e-heritage.ru/>
- Интелрос. Интеллектуальная Россия: <http://www.intelros.ru/>
- Национальная электронная библиотека НЭБ: <http://www.rusneb.ru>
- Неприкосновенный запас: <http://magazines.russ.ru/nz/>
- Президентская библиотека: <http://www.prlib.ru>
- Российская государственная библиотека: <http://www.rsl.ru/>
- Российская национальная библиотека: <http://www.nlr.ru/poisk/>

7. ПРОГРАММНОЕ И МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ПРОГРАММЫ

В ходе реализации образовательного процесса используются многофункциональные аудитории для проведения занятий лекционного типа, занятий семинарского типа, укомплектованные специализированной мебелью и техническими средствами обучения, служащими для представления учебной информации большой аудитории.

Проведение занятий лекционного типа и семинарского типа обеспечивается демонстрационным оборудованием.

Для обучающихся с ограниченными возможностями здоровья и инвалидов (в случае необходимости) могут быть созданы специальные условия для получения образования.

Программное обеспечение

При осуществлении образовательного процесса в рамках Университета слушателям рекомендовано использовать следующее лицензионное программное обеспечение:

- OS Microsoft Windows (OVS OS Platform)
- MS Office (OVS Office Platform)
- Adobe Acrobat Professional 11.0 MLP AOO License RU
- Adobe CS5.5 Design Standart Win IE EDU CLP
- ABBYY FineReader 11 Corporate Edition
- ABBYY Lingvo x5
- Adobe Photoshop Extended CS6 13.0 MLP AOO License RU
- Adobe Acrobat Reader DC /Pro – бесплатно
- Google Chrome – бесплатно
- Opera – бесплатно
- Mozilla – бесплатно
- Яндекс.Браузер – бесплатно
- VLC – бесплатно