

Министерство науки и высшего образования Российской Федерации
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ (НИ ТГУ)

Институт прикладной математики и компьютерных наук

УТВЕРЖДАЮ

Директор института прикладной
математики и компьютерных наук

А.В. Замятин

« 18 » 03 2022 г.

Рабочая программа дисциплины

Методы машинного обучения в телекоммуникациях

по направлению подготовки

**02.04.02 Фундаментальная информатика и
информационные технологии**

Профиль подготовки:

«Математика беспроводных сетей связи и интернета вещей»

Форма обучения

Очная

Квалификация

Магистр

Год приема

2022

Код дисциплины в учебном плане: ФТД.02

СОГЛАСОВАНО:

Руководитель ОП

С.П. Моисеева

Председатель УМК

С.П. Сущенко

Томск – 2022

1. Цель и планируемые результаты освоения дисциплины (модуля)

Одним из характерных особенностей конца XX и начала XXI веков является бурное развитие и увеличение числа сетей передачи информации. В ближайшие десятилетия эта тенденция будет только нарастать в свете глобальной цифровизации, охватившей большинство стран мира. Как локальные сети, которые охватывают промышленность, сферу производства, образование, медицину и другие отрасли, так и распределенные сети, связывающие города, регионы и континенты, все больше проникают во все сферы человеческой деятельности.

Современные сети связи предоставляют широкий набор услуг передачи текстовой, аудио- и видеоданных, работа с удаленными базами данных, сенсорных сетей и др. в реальном времени. В настоящее время сети передачи данных обеспечивают дистанционное обучение, телемедицину, телеконференции, торговлю в Интернет-магазинах, поисковые запросы, удаленный режим работы компаний, и т. д. Все это сопровождается разработкой новых сетевых технологий, способствующих повышению скорости, приоритетности и надежности передачи.

В рамках дисциплины происходит обучение аппарату машинного обучения для анализа и количественной оценки процессов обслуживания информационных потоков в системах распределения информации (системах телетрафика) в том числе для описания процессов передачи данных в IoT, БПЛА, а также смежных сквозных технологий.

Целью освоения дисциплины является формирование следующих компетенций: – ОПК-3. Способность проводить анализ математических моделей, создавать инновационные методы решения прикладных задач профессиональной деятельности в области информатики и математического моделирования

Результатами освоения дисциплины являются следующие индикаторы достижения компетенций:

ИОПК-3.1 Проводит анализ математических моделей и систем

ИОПК-3.2 Применяет математические модели, методы для решения прикладных задач профессиональной деятельности

2. Задачи освоения дисциплины

1. Изучить понятия и методы машинного обучения.
2. Изучить наиболее популярные направления исследований в машинном обучении.
3. Освоить теоретические, математические и алгоритмические основы реализации технологии машинного обучения для решения задач в области телекоммуникаций.

3. Место дисциплины (модуля) в структуре образовательной программы

Дисциплина «Методы машинного обучения в телекоммуникациях» относится к факультативным дисциплинам.

4. Семестр(ы) освоения и форма(ы) промежуточной аттестации по дисциплине

Семестр 2, зачет .

5. Входные требования для освоения дисциплины

Для успешного освоения дисциплины требуются компетенции, сформированные в ходе освоения образовательных программ предшествующего уровня образования.

Для успешного освоения дисциплины требуются результаты обучения по следующим дисциплинам: Теория телетрафика, Интернет вещей, Математические модели телекоммуникационных потоков.

6. Язык реализации

Русский.

7. Объем дисциплины (модуля)

Общая трудоемкость дисциплины составляет 2 з.е., 72 часа, из которых:

– лекции: 16 ч.;

– семинарские занятия: 0 ч.

– практические занятия: 0 ч.;

– лабораторные работы: 16 ч.

в том числе практическая подготовка: 0 ч.

Объем самостоятельной работы студента определен учебным планом.

8. Содержание дисциплины (модуля), структурированное по темам

Тема 1. Введение в машинное обучение
Постановки основных классов задач в машинном обучении. Обучение с учителем (supervised learning): регрессия и классификация; обучение без учителя (unsupervised learning): кластеризация, снижение размерности; semi-supervised learning, рекомендательные системы, обработка текстов: тематическое моделирование, построение аннотаций, извлечение ответов на вопросы, машинный перевод; обработка изображений: порождение, преобразование; обучение представлений; обучение с подкреплением. Примеры задач. Виды данных: структурированные таблицы, тексты, изображения, звук, логи. Признаки.

Тема 2. Признаковые представления для дискретных входных данных
Кодирование категориальных данных. Пропущенные значения. Обработка текстов: bag of words, tf-idf, векторные эмбединги.

Тема 3. Выбор и оценка моделей, работа с признаками
Кросс-валидация (отбор переменных, переобучение на валидационное множество). Оценки ожидаемой ошибки для линейной регрессии: AIC и другие. L1 и L2 регуляризация. Методы отбора признаков. Метод главных компонент и singular spectrum analysis. Ядровые методы. Ядра и спрямляющие пространства, методы их построения. Операции в спрямляющих пространствах.

Тема 4. Линейные модели и задача классификации
Задачи классификации. Общая постановка. 0-1 ошибка. Байесовский классификатор. Линейные методы для классификации. Логистическая регрессия, максимизация правдоподобия, кросс-энтропия.

Тема 5. Линейные методы регрессии и классификации
Постановка задач обучения. Объекты и признаки, вектора признаков, классы. Типы задач: классификация, регрессия, кластеризация, ранжирование. Основные понятия: методы

обучения, функция потерь и функционал качества, переобучение, обобщающая способность. Примеры прикладных задач.

Тема 6. Введение в линейные модели и задача регрессии *Градиентный спуск, методы оценивания градиента. Функции потерь. Метрики качества регрессии. Линейная регрессия, метод наименьших квадратов и максимизация правдоподобия. Теорема Гаусса—Маркова. Явный вид решения в методе наименьших квадратов. Ковариационная матрица для коэффициентов.*

Тема 7. Машинное обучение как математическое моделирование *Статистические модели. Теоретико-вероятностная постановка задачи обучения с учителем. Минимизация ожидаемой ошибки. No free lunch theorem. Пример: задача регрессии, минимизация квадрата отклонения. Регрессионная функция. Линейная регрессия и метод k ближайших соседей. Переобучение и недообучение.*

9. Текущий контроль по дисциплине

Текущий контроль по дисциплине проводится путем контроля посещаемости, проведения лабораторных работ и фиксируется в форме контрольной точки не менее одного раза в семестр.

10. Порядок проведения и критерии оценивания промежуточной аттестации

Зачет ставится по результатам ответов на вопросы, включающих в себя:

- Байесовский классификатор.
- Линейные методы для классификации.
- Логистическая регрессия.
- Максимизация правдоподобия.
- Кросс-энтропия.

Результаты зачета определяются оценками «зачтено», «незачтено».

11. Учебно-методическое обеспечение

а) Электронный учебный курс по дисциплине в электронном университете «Moodle» - <https://moodle.tsu.ru/course/view.php?id=00000>

б) Оценочные материалы текущего контроля и промежуточной аттестации по дисциплине.

12. Перечень учебной литературы и ресурсов сети Интернет

а) основная литература:

1. Флах П. Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных - Издательство "ДМК Пресс" - 2015 - 400с. - ISBN: 978-5-97060-273-7 - Текст электронный // ЭБС ЛАНЬ - URL: <https://e.lanbook.com/book/69955>
2. Мерков А. Б. Распознавание образов. Введение в методы статистического обучения. 2011. 256 с.
3. Вьюгин В.В. Математические основы машинного обучения и прогнозирования - Московский центр непрерывного математического образования - 2014 - 304с. - ISBN: 978-5-4439-2014-6 - Текст электронный // ЭБС ЛАНЬ - URL: <https://e.lanbook.com/book/5639>

б) дополнительная литература:

1. Коэльо Л.П., Ричарт В. - Построение систем машинного обучения на языке Python - Издательство "ДМК Пресс" - 2016 - 302с. - ISBN: 978-5-97060-330-7 - Текст электронный // ЭБС ЛАНЬ - URL: <https://e.lanbook.com/book/82818>

2. Рашка С. Python и машинное обучение: крайне необходимое пособие по новейшей предсказательной аналитике, обязательное для более глубокого понимания методологии машинного обучения - Издательство "ДМК Пресс" - 2017 - 418с. - ISBN: 978-5-97060-409-0 - Текст электронный // ЭБС ЛАНЬ - URL: <https://e.lanbook.com/book/100905>

в) ресурсы сети Интернет:

– Научная электронная библиотека – <https://www.elibrary.ru/>

–Международные научные базы цитирования <https://www.scopus.com/> и <https://www.webofknowledge.com/>

Тематические научные журналы:

– Machine Learning <https://www.springer.com/journal/10994>

– Mathematics (<https://www.mdpi.com/journal/mathematics>)

13. Перечень информационных технологий

а) лицензионное и свободно распространяемое программное обеспечение:

– Microsoft Office Standart 2013 Russian: пакет программ. Включает приложения: MS Office Word, MS Office Excel, MS Office PowerPoint, MS Office On-eNote, MS Office Publisher, MS Outlook, MS Office Web Apps (Word Excel MS PowerPoint Outlook);

– математический пакет программ MathCad

– публично доступные облачные технологии (Google Docs, Яндекс диск и т.п.).

– инструменты видео-конференций (Adobe Connect, Яндекс.Телемост, Zoom, Voov)

– онлайн-доски Jamboard, Miro, SBoard,

– вспомогательные цифровые инструменты для образовательного процесса (Mentimeter, Yandex Forms, Overleaf, PDF-XChange Viewer, Яндекс.Контекст)

б) информационные справочные системы:

– Электронный каталог Научной библиотеки ТГУ – <http://chamo.lib.tsu.ru/search/query?locale=ru&theme=system>

– Электронная библиотека (репозиторий) ТГУ – <http://vital.lib.tsu.ru/vital/access/manager/Index>

– ЭБС Лань – <http://e.lanbook.com/>

– ЭБС Консультант студента – <http://www.studentlibrary.ru/>

– Образовательная платформа Юрайт – <https://urait.ru/>

– ЭБС ZNANIUM.com – <https://znanium.com/>

– ЭБС IPRbooks – <http://www.iprbookshop.ru/>

14. Материально-техническое обеспечение

Аудитории для проведения занятий лекционного типа.

Аудитории для проведения занятий семинарского типа, индивидуальных и групповых консультаций, текущего контроля и промежуточной аттестации.

Помещения для самостоятельной работы, оснащенные компьютерной техникой и доступом к сети Интернет, в электронную информационно-образовательную среду и к информационным справочным системам.

Лаборатории, оборудованные ПК с лицензионной ПО математического пакета MathCad

15. Информация о разработчиках

Бронер Валентина Игоревна - кандидат физ.-мат. наук, доцент кафедры теории вероятностей и математической статистики ИПМКН ТГУ