

**Федеральное государственное автономное образовательное
учреждение высшего образования
«Московский физико-технический институт
(национальный исследовательский университет)»**

УТВЕРЖДЕНО
Заместитель директора

Ю.О. Соболев

	Рабочая программа дисциплины (модуля)
по дисциплине:	Соревновательный DS
по направлению:	Прикладная математика и информатика
профиль подготовки:	Науки о данных центр дополнительного, дополнительного профессионального и онлайн-образования "Пуск" центр дополнительного, дополнительного профессионального и онлайн-образования "Пуск"
курс:	2
квалификация:	магистр

Семестр, формы промежуточной аттестации: 4 (весенний) - Дифференцированный зачет

Аудиторных часов: 22 всего, в том числе:

лекции: 2 час.

семинары: 20 час.

лабораторные занятия: 0 час.

Самостоятельная работа: 113 час.

Всего часов: 135, всего зач. ед.: 3

Программу составили:

А.А. Миленькин, преподаватель

К.А. Лапин, старший методист

Программа обсуждена на заседании центра дополнительного, дополнительного профессионального и
онлайн-образования "Пуск" 01.03.2025

Аннотация

Студенты получают знания о особенностях применения методов машинного обучения в соревновательных задачах по ds

1. Цели и задачи

Цель дисциплины

- обучить студентов особенностям применения методов машинного обучения соревновательных задачах по ds.

Задачи дисциплины

- дать студентам нужные практические навыки в применении методов машинного обучения в соревновательных задачах по ds.

2. Перечень формируемых компетенций

Освоение дисциплины направлено на формирование следующих компетенций:

Код и наименование компетенции	Индикаторы достижения компетенции
УК-1 Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий	УК-1.1 Анализирует проблемную ситуацию как систему, выявляя ее составляющие и связи между ними
	УК-1.2 Осуществляет поиск вариантов решения поставленной проблемной ситуации на основе доступных источников информации
	УК-1.3 Разрабатывает стратегию достижения поставленной цели как последовательность шагов, предвидя результат каждого из них и оценивая их влияние на внешнее окружение планируемой деятельности и на взаимоотношения участников этой деятельности
ОПК-2 Способен совершенствовать и реализовывать новые математические методы решения прикладных задач	ОПК-2.1 Имеет представление о современном состоянии математических исследований в рамках тематической области своей профессиональной деятельности
	ОПК-2.2 Способен оценить актуальность и практическую значимость прикладных математических исследований в своей профессиональной области
ОПК-5 Имеет представление об актуальных проблемах науки и техники в области информатики и вычислительной техники, способен на научном языке формулировать профессиональные задачи	ОПК-5.1 Имеет представление о современном состоянии исследований в рамках тематической области своей профессиональной деятельности
	ОПК-5.2 Способен оценивать актуальность исследований в области информатики и вычислительной техники и их практическую значимость
	ОПК-5.3 Владеет профессиональной терминологией, используемой в современной научно-технической литературе, обладает навыками устного и письменного изложения результатов научной деятельности в рамках профессиональной коммуникации
ОПК-6 Способен выбирать и (или) разрабатывать подходы к решению типовых и новых задач в области информатики и	ОПК-6.1 Способен анализировать задачу, планировать пути решения, предлагать и комбинировать способы решения
	ОПК-6.2 Способен разрабатывать и модернизировать программное и аппаратное обеспечение информационных и автоматизированных систем
	ОПК-6.3 Способен использовать исследовательские методы при решении новых задач, применяя знания из различных областей науки (техники)

вычислительной техники, учитывая особенности и ограничения различных методов решения	ОПК-6.5 Способен самостоятельно приобретать, развивать и применять математические, естественнонаучные, социально-экономические и профессиональные знания для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте
	ОПК-6.4 Владеет аналитическими и вычислительными методами решения, понимает и учитывает на практике границы применимости получаемых решений

3. Перечень планируемых результатов обучения по дисциплине (модулю)

В результате освоения дисциплины обучающиеся должны

знать:

- специфику соревновательных задач;
- текущие state of the art методы.

уметь:

- строить самостоятельно решения для задач по ds;
- применять различные state of the art методы для увел. производительности. моделей;
- предобрабатывать данные для для увел. производительности. моделей;
- строить схемы валидации моделей.

владеть:

- навыками и инструментами необходимыми для успешного решения задач в хакатонах по ds.

4. Содержание дисциплины (модуля), структурированное по темам (разделам) с указанием отведенного на них количества академических часов и видов учебных занятий

4.1. Разделы дисциплины (модуля) и трудоемкости по видам учебных занятий

№	Тема (раздел) дисциплины	Трудоемкость по видам учебных занятий, включая самостоятельную работу, час.			
		Лекции	Семинары	Лаборат. работы	Самост. работа
1	Генерация признаков на основе имеющихся данных	1	4		20
2	Обучение и тюнинг ML моделей	1	4		20
3	Фреймворки для оптимизации работы моделей		5		32
4	Ансамблирование нескольких решений		3		22
5	Вспомогательные техники: оптимизация памяти, ускорение вычислений, парсинг внешних данных		4		19
Итого часов		2	20		113
Подготовка к экзамену		0 час.			
Общая трудоёмкость		135 час., 3 зач.ед.			

4.2. Содержание дисциплины (модуля), структурированное по темам (разделам)

Семестр: 4 (Весенний)

1. Генерация признаков на основе имеющихся данных

Feature Engineering. Feature Selection. Продвинутая визуализация

2. Обучение и тюнинг ML моделей

Catboost. LightGBM. XgBoost

3. Фреймворки для оптимизации работы моделей

Автоматическая генерация и фильтрация признаков. Автоматическая оптимизация ML моделей.

4. Ансамблирование нескольких решений

Автоматическая генерация и фильтрация признаков. Автоматическая оптимизация ML моделей

5. Вспомогательные техники: оптимизация памяти, ускорение вычислений, парсинг внешних данных

Оптимизация памяти и ускорение вычислений. Парсинг внешних данных.

5. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине (модулю)

В обучение используются дистанционные образовательные технологии.

6.Перечень рекомендуемой литературы

Основная литература

1. Динамическое программирование в примерах и задачах [Текст] : учеб. пособие: доп. М-вом высш. и сред. спец. образования СССР / И. Л. Калихман, М. А. Войтенко .— М : Высшая школа, 1979 .— 125 с.
2. Алгоритмы обработки экспериментальных данных [Текст]/отв. ред. И. А. Овсевич , -М., Наука, 1986

Дополнительная литература

7. Перечень ресурсов информационно-телекоммуникационной сети "Интернет", необходимых для освоения дисциплины (модуля)

Не используются

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

1. Google Drive для доступа к материалам курса
2. Zoom
3. Ноутбук для участия в интерактивных занятиях.

9. Методические указания для обучающихся по освоению дисциплины (модуля)

Студенту рекомендуется внимательно слушать лектора, следить за тем, что написано на доске или представлено на слайдах презентации, анализировать получаемую им информацию. В случае, если материал лекции непонятен, следует задать вопрос в отведенное для вопросов время. Студенту также рекомендуется конспектировать материал лекции в тетради, что улучшает запоминание.

При выполнении практических работ студенту рекомендуется внимательно анализировать поставленную задачу, уделяя особое внимание критериям оценки точности решения задачи. Особенное внимание следует уделять методологическим аспектам решения задач.

При ведении самостоятельной работы студенту рекомендуется внимательно подходить к изучению научных статей, обращать внимание на значимость полученного результата, на требования к обучающей выборке, на скорость работы предлагаемых алгоритмов, на результаты их сравнения с существующими. В случае, если изучаемый материал понятен не до конца, рекомендуется обращение к дополнительной литературе.

ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

по направлению:	Прикладная математика и информатика
профиль подготовки:	Науки о данных центр дополнительного, дополнительного профессионального и онлайн-образования "Пуск" центр дополнительного, дополнительного профессионального и онлайн-образования "Пуск"
курс:	<u>2</u>
квалификация:	магистр

Семестр, формы промежуточной аттестации: 4 (весенний) - Дифференцированный зачет

Разработчики:

А.А. Миленькин, преподаватель

К.А. Лапин, старший методист

1. Компетенции, формируемые в процессе изучения дисциплины

Код и наименование компетенции	Индикаторы достижения компетенции
УК-1 Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий	УК-1.1 Анализирует проблемную ситуацию как систему, выявляя ее составляющие и связи между ними
	УК-1.2 Осуществляет поиск вариантов решения поставленной проблемной ситуации на основе доступных источников информации
	УК-1.3 Разрабатывает стратегию достижения поставленной цели как последовательность шагов, предвидя результат каждого из них и оценивая их влияние на внешнее окружение планируемой деятельности и на взаимоотношения участников этой деятельности
ОПК-2 Способен совершенствовать и реализовывать новые математические методы решения прикладных задач	ОПК-2.1 Имеет представление о современном состоянии математических исследований в рамках тематической области своей профессиональной деятельности
	ОПК-2.2 Способен оценить актуальность и практическую значимость прикладных математических исследований в своей профессиональной области
ОПК-5 Имеет представление об актуальных проблемах науки и техники в области информатики и вычислительной техники, способен на научном языке формулировать профессиональные задачи	ОПК-5.1 Имеет представление о современном состоянии исследований в рамках тематической области своей профессиональной деятельности
	ОПК-5.2 Способен оценивать актуальность исследований в области информатики и вычислительной техники и их практическую значимость
	ОПК-5.3 Владеет профессиональной терминологией, используемой в современной научно-технической литературе, обладает навыками устного и письменного изложения результатов научной деятельности в рамках профессиональной коммуникации
ОПК-6 Способен выбирать и (или) разрабатывать подходы к решению типовых и новых задач в области информатики и вычислительной техники, учитывая особенности и ограничения различных методов решения	ОПК-6.1 Способен анализировать задачу, планировать пути решения, предлагать и комбинировать способы решения
	ОПК-6.2 Способен разрабатывать и модернизировать программное и аппаратное обеспечение информационных и автоматизированных систем
	ОПК-6.3 Способен использовать исследовательские методы при решении новых задач, применяя знания из различных областей науки (техники)
	ОПК-6.5 Способен самостоятельно приобретать, развивать и применять математические, естественнонаучные, социально-экономические и профессиональные знания для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте
	ОПК-6.4 Владеет аналитическими и вычислительными методами решения, понимает и учитывает на практике границы применимости получаемых решений

2. Показатели оценивания компетенций

В результате изучения дисциплины «Соревновательный DS» обучающийся должен:

знать:

- специфику соревновательных задач;
- текущие state of the art методы.

уметь:

- строить самостоятельно решения для задач по ds;
- применять различные state of the art методы для увел. производительности. моделей;
- предобрабатывать данные для для увел. производительности. моделей;
- строить схемы валидации моделей.

владеть:

- навыками и инструментами необходимыми для успешного решения задач в хакатонах по ds.

3. Перечень типовых (примерных) вопросов, заданий, тем для подготовки к текущему контролю

Примеры вопросов для текущего контроля:

1. В чем особенность использования алгоритма оптимизации Adam
2. Продвинутое техники предобработки данных.
3. Продвинутое техники аугментаций данных
4. Продвинутое техники генераций новых признаков
5. Продвинутое библиотеки для оптимизации гиперпараметров
6. Продвинутое техники регуляризации моделей

4. Перечень типовых (примерных) вопросов и тем для проведения промежуточной аттестации обучающихся

Примерные вопросы для дифференцированного зачета:

1. Методы аугментации данных
2. Алгоритм Stochastic weight averaging
3. Очистка данных используется для обнаружения, исправления или удаления ошибочных записей в наборе данных[4];
4. Нормализация данных.
5. Преобразование данных.
6. Выделение признаков.
7. Уплотнение данных.

Критерии оценивания

- оценка «отлично (10)» выставляется студенту, показавшему всесторонние, систематизированные, глубокие знания учебной программы дисциплины и умение уверенно применять их на практике при решении конкретных задач, свободное и правильное обоснование принятых решений
- оценка «отлично (9)» выставляется студенту, показавшему всесторонние, систематизированные, глубокие знания учебной программы дисциплины и умение применять их на практике при решении конкретных задач, свободное и правильное обоснование принятых решений
- оценка «отлично (8)» выставляется студенту, показавшему всесторонние систематизированные, глубокие знания учебной программы дисциплины и умение применять их на практике при решении конкретных задач, и правильное обоснование принятых решений
- оценка «хорошо (7)» выставляется студенту, если он твердо знает материал, грамотно и, по существу, излагает его, умеет применять полученные знания на практике, но допускает в ответе или в решении задач некоторые неточности;
- оценка «хорошо (6)» выставляется студенту, если он знает материал, грамотно и, по существу, излагает его, умеет применять полученные знания на практике, но допускает в ответе или в решении задач некоторые неточности;
- оценка «хорошо (5)» выставляется студенту, если он знает материал, и, по существу, излагает его, умеет применять полученные знания на практике, но допускает в ответе или в решении задач некоторые неточности;

- оценка «удовлетворительно (4)» выставляется студенту, показавшему фрагментарный, разрозненный характер знаний, недостаточно правильные формулировки базовых понятий, нарушения логической последовательности в изложении программного материала, но при этом он владеет основными разделами учебной программы, необходимыми для дальнейшего обучения и может применять полученные знания по образцу в стандартной ситуации;
- оценка «удовлетворительно (3)» выставляется студенту, показавшему фрагментарный, разрозненный характер знаний, недостаточно правильные формулировки базовых понятий, нарушения логической последовательности в изложении программного материала, но при этом он владеет фрагментарно основными разделами учебной программы, необходимыми для дальнейшего обучения и может применять полученные знания по образцу в стандартной ситуации;
- оценка «неудовлетворительно (2)» выставляется студенту, который не знает большей части основного содержания учебной программы дисциплины, допускает грубые ошибки в формулировках основных понятий дисциплины и не умеет использовать полученные знания при решении типовых практических задач
- оценка «неудовлетворительно (1)» выставляется студенту, который не знает формулировок основных понятий дисциплины.

5. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности

Итоговый контроль проводится в формате дифференцированного зачета на lms.

На дифференцированный зачет отводится 2 академических часа