



МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Магнитогорский государственный технический университет им. Г.И. Носова»



УТВЕРЖДАЮ
Директор ИИиС
И.Ю. Мезин

30.01.2023 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

***ФОРМИРОВАНИЕ ОБУЧАЮЩИХ НАБОРОВ ДАННЫХ В
МЕТАЛЛУРГИИ***

Направление подготовки (специальность)
22.04.02 Metallurgy

Направленность (профиль/специализация) программы
Искусственный интеллект в металлургии

Уровень высшего образования - магистратура

Форма обучения
очная

Институт/ факультет	Институт естествознания и стандартизации
Кафедра	Прикладной математики и информатики
Курс	1
Семестр	2

Магнитогорск
2023 год

Рабочая программа составлена на основе ФГОС ВО - магистратура по направлению подготовки 22.04.02 Metallургия (приказ Минобрнауки России от 24.04.2018 г. № 308)

Рабочая программа рассмотрена и одобрена на заседании кафедры Прикладной математики и информатики
17.01.2023, протокол № 5

Зав. кафедрой  Ю.А. Извеков

Рабочая программа одобрена методической комиссией ИЕиС
30.01.2023 г. протокол № 5

Председатель  И.Ю. Мезин

Согласовано:
Зав. кафедрой Литейных процессов и материаловедения

 Н.А. Феоктистов

Рабочая программа составлена:
доцент кафедры ПМИИ, канд. пед. наук  Л.С. Рязанова

Рецензент:
доцент кафедры Физики, канд. физ.-мат. наук  Д.М.
Долгушин

Лист актуализации рабочей программы

Рабочая программа пересмотрена, обсуждена и одобрена для реализации в 2024 - 2025 учебном году на заседании кафедры Прикладной математики и информатики

Протокол от _____ 20__ г. № ____
Зав. кафедрой _____ Ю.А. Извеков

Рабочая программа пересмотрена, обсуждена и одобрена для реализации в 2025 - 2026 учебном году на заседании кафедры Прикладной математики и информатики

Протокол от _____ 20__ г. № ____
Зав. кафедрой _____ Ю.А. Извеков

1 Цели освоения дисциплины (модуля)

Сформировать у студентов навыки работы с данными и решения прикладных задач, дать представление о искусственном интеллекте, об основных методах машинного обучения и видах задач, решаемых ими. Ознакомить обучающихся со способами формирования обучающих наборов данных.

2 Место дисциплины (модуля) в структуре образовательной программы

Дисциплина Формирование обучающих наборов данных в металлургии входит в часть учебного плана формируемую участниками образовательных отношений образовательной программы.

Для изучения дисциплины необходимы знания (умения, владения), сформированные в результате изучения дисциплин/ практик:

Искусственный интеллект и машинное обучение

Знания (умения, владения), полученные при изучении данной дисциплины будут необходимы для изучения дисциплин/практик:

Контроль технологических процессов на металлургических предприятиях с использованием искусственного интеллекта

Проектирование технологических процессов с использованием искусственного интеллекта

3 Компетенции обучающегося, формируемые в результате освоения дисциплины (модуля) и планируемые результаты обучения

В результате освоения дисциплины (модуля) «Формирование обучающих наборов данных в металлургии» обучающийся должен обладать следующими компетенциями:

Код индикатора	Индикатор достижения компетенции
ПК-8	Способен адаптировать и применять методы и алгоритмы машинного обучения для решения прикладных задач в различных предметных областях
ПК-8.1	Ставит задачи по адаптации или совершенствованию методов и алгоритмов для решения комплекса задач предметной области Знает: классы методов и алгоритмов машинного обучения; классы методов и алгоритмов машинного обучения Умеет: ставить задачи и адаптировать методы и алгоритмы машинного обучения Имеет практический опыт: участия в проектах по изучению опыта адаптации и применимости методов и алгоритмов машинного обучения для решения прикладных задач в металлургии; постановки задач по адаптации или совершенствованию методов и алгоритмов для решения комплекса задач предметной области

4. Структура, объём и содержание дисциплины (модуля)

Общая трудоемкость дисциплины составляет 4 зачетных единиц 144 академических часов, в том числе:

- контактная работа – 45 академических часов;
- аудиторная – 45 академических часов;
- внеаудиторная – 0 академических часов;
- самостоятельная работа – 99 академических часов;
- в форме практической подготовки – 0 академических часов;

Форма аттестации - курсовая работа, зачет

Раздел/ тема дисциплины	Семестр	Аудиторная контактная работа (в очной форме)			Самостоятельная работа студента	Вид самостоятельной работы	Форма текущего контроля успеваемости и промежуточной аттестации	Код компетенции
		Лек.	лаб. зан.	практ. зан.				
1. Способы формирования								
1.1 Понятие DataSet и его	2	2			4	Изучение литературы	Опрос, беседа	ПК-8.1
1.2 Поиск и структуризация открытых данных		2		4	12	Изучение литературы. Выполнение лабораторной работы	Проверка выполнения лабораторной работы	ПК-8.1
1.3 Классификация задач, решаемых с помощью машинного		2		10	12	Изучение литературы. Выполнение лабораторной работы	Проверка выполнения лабораторной работы	ПК-8.1
1.4 Формирование обучающих выборок для		5		4	30	Изучение литературы. Выполнение лабораторной работы	Проверка выполнения лабораторной работы	ПК-8.1
1.5 Формирование обучающих выборок для		4		12	41	Изучение литературы. Выполнение лабораторной работы	Проверка выполнения лабораторной работы	ПК-8.1
Итого по разделу		15		30	99			
Итого за семестр		15		30	99		зачёт,кр	
Итого по дисциплине		15		30	99		курсовая работа, зачет	

5 Образовательные технологии

В ходе изучения дисциплины рекомендуется использовать образовательные и информационные технологии:

1. Традиционные технологии обучения, предполагающие передачу информации в готовом виде, формирование учебных умений по образцу: лекция-изложение, лекция-объяснение, лабораторные работы, контрольная работа и др.

Использование традиционных технологий обеспечивает ориентирование студента в потоке информации, связанной с различными подходами к определению сущности, содержания, методов, форм развития и саморазвития личности; самоопределение в выборе оптимального пути и способов личностно-профессионального развития; систематизацию знаний, полученных студентами в процессе аудиторной и самостоятельной работы. Лекционные занятия проводятся с использованием мультимедийных средств. Лабораторные занятия обеспечивают развитие и закрепление умений и навыков определения целей и задач саморазвития, а также принятия наиболее эффективных решений по их реализации.

2. Информационно-коммуникационные образовательные технологии – организация образовательного процесса, основанная на применении специализированных программных сред и технических средств работы с информацией.

Лабораторные занятия проводятся в компьютерных классах вычислительного центра ФГБОУ ВО «МГТУ». В ходе проведения лекционных занятий предусматривается: использование образовательного портала ФГБОУ ВО «МГТУ».

3. Интерактивные формы обучения, предполагающие организацию обучения как продуктивной творческой деятельности в режиме взаимодействия студентов друг с другом и с преподавателем. Использование интерактивных образовательных технологий способствует повышению интереса и мотивации учащихся, активизации мыслительной деятельности и творческого потенциала студентов, делает более эффективным усвоение материала, позволяет индивидуализировать обучение и ввести экстренную коррекцию знаний.

4. Проблемная технология обучения

Методика, ориентирована на лекции проблемно-информационного характера, лабораторные работы поисково-исследовательского типа

Проблемная лекция

Часть теоретического материала изучается в форме традиционных лекций, другая часть проблемным методом. Начиная с создания познавательной потребности в решении возникшей в результате постановки учебной проблемной ситуации, необходимо добиться осознания студентами проблемы, провести поиск гипотезы, касающейся результата и пути его получения. Часть лекционного материала излагается с использованием информационных технологий (средств мультимедиа), что способствует познавательной активности обучающихся. Решение проблемы, которое является основой перехода к следующей учебной проблеме и ведет к открытию нового знания, предполагается проводить вне лекционных часов.

Лабораторные работы поисково-исследовательского типа

Обмен информацией, полученной студентами в ходе самостоятельного поиска и исследования по поставленной проблеме, рекомендуется организовать в рамках лабораторных работ. Ценность данной формы занятий в том, что в процессе обсуждения можно высказать собственное мнение и попытаться доказать его правильность.

При изучении дисциплины для каждого раздела предлагается перечень вопросов для самоконтроля. Возможны три варианта использования данных вопросов при изучении теоретического материала: либо для контроля полученных студентами знаний по окончании изучения раздела, либо для обсуждения каждого вопроса как

мини-проблемы в ходе лабораторной работы, либо то и другое в определенном сочетании. До-пускается иная постановка вопросов преподавателем, а самостоятельная формулировка студентами вопросов для обсуждения при выполнении лабораторной работы только приветствуется. Лабораторные работы поисково-исследовательского типа не только способствует углубленной проработке теоретического материала предмета на протяжении всего изучения курса, но и развивают творческую самостоятельность студентов, способность к обобщениям, укрепляя их интерес к исследованиям, содействуя выработке практических навыков работы.

6 Учебно-методическое обеспечение самостоятельной работы обучающихся
Представлено в приложении 1.

7 Оценочные средства для проведения промежуточной аттестации
Представлены в приложении 2.

8 Учебно-методическое и информационное обеспечение дисциплины (модуля)

а) Основная литература:

1. Обучение нейронной сети. [Электронный ресурс]. Режим доступа: <https://www.bigdataschool.ru/bigdata/dataset-data-preparation.html/> (дата обращения: 23.12.2019).

1. Лимановская, О. В. Основы машинного обучения: учебное пособие / О.В. Лимановская, Т. И. Алферьева. - 2-е изд., стер. - Москва: ФЛИНТА: Изд-во Урал. ун-та, 2022. - 88 с

б) Дополнительная литература:

1. Регрессионный анализ. Дисперсионный анализ. [Электрон. ресурс].Режим доступа: Microsoft PowerPoint - Лекция 3 (tri.ru). Дата обращения:04.05.2023.

2. М.А. Поручиков. Анализ данных: учеб. пособие / М.А. Поручиков. –Самара: Изд-во Самарского университета, 2016. – 88 с.

в) Методические указания:

1.Соколов, Г.А. Введение в регрессионный анализ и планирование регрессионных экспериментов: Учебное пособие / Г.А. Соколов, Р.В. Сагитов. - М.: Инфра-М, 2016. - 352 с.

2. Кулаичев, А.П. Методы и средства комплексного анализа данных: Учебное пособие / А.П. Кулаичев. — М.: Форум, НИЦ ИНФРА-М, 2016. — 512 с.

г) Программное обеспечение и Интернет-ресурсы:

Программное обеспечение

Наименование ПО	№ договора	Срок действия лицензии
STATISTICA в.6	К-139-08 от 22.12.2008	бессрочно
7Zip	свободно распространяемое ПО	бессрочно
MS Office 2007 Professional	№ 135 от 17.09.2007	бессрочно

Профессиональные базы данных и информационные справочные системы

Название курса	Ссылка

9 Материально-техническое обеспечение дисциплины (модуля)

Материально-техническое обеспечение дисциплины включает:

Материально-техническое обеспечение дисциплины включает:

1) Учебные аудитории для проведения занятий лекционного типа: Доска, мультимедийный проектор, экран

2) Учебные аудитории для проведения лабораторных занятий, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации: Персональные компьютеры с пакетом MS Office, выходом в Интернет и с доступом в электронную информационнообразовательную среду университета; Комплекс тестовых заданий для проведения рубежного и промежуточного контроля.

3) Помещения для самостоятельной работы обучающихся: Персональные компьютеры с пакетом MS Office, выходом в Интернет и с доступом в электронную информационно-образовательную среду университета.

4) Помещения для хранения и профилактического обслуживания учебного оборудования: Шкафы для хранения учебно-методической документации, учебного оборудования и учебно-наглядных пособий

Приложение 1. Учебно-методическое обеспечение самостоятельной работы обучающихся

Вопросы для самоконтроля

- 1 Приведите примеры непрерывных данных.
- 2 Приведите примеры категориальных данных.
- 3 Дайте определения понятию «источник данных».
- 4 Приведите способы классификации источников данных.
- 5 Охарактеризуйте понятие «открытые данные».
- 6 Приведите примеры источников открытых данных.
- 7 Перечислите основные форматы хранения данных.
- 8 Приведите алгоритм построения системы сбора данных на основе программного обеспечения Microsoft Excel.
- 9 Обоснуйте необходимость подготовки данных.
- 10 Охарактеризуйте операцию форматирования данных.
- 11 Приведите пример форматирования данных.
- 12 Охарактеризуйте операцию отбора данных.
- 13 Приведите пример отбора данных
- 14 Охарактеризуйте операцию нормализации данных.
- 15 Приведите пример нормализации данных.
- 16 Охарактеризуйте операцию кодирования данных.
- 17 Приведите пример кодирования данных
- 18 Охарактеризуйте понятие классификации данных.
- 19 Виды классификации данных.
- 20 Перечислите методы классификации данных.
- 21 Приведите пример практического применения классификации.
- 22 Приведите алгоритм применения логистической регрессии.
- 23 Запишите функции штрафа при логистической регрессии.
- 24 Поясните суть проблемы линейного разделения классов.
- 25 Дайте определение ошибки классификации первого рода.
- 26 Дайте определение ошибки классификации второго рода.

- 27 Приведите примеры ошибок классификации и возможных последствий.
- 28 Приведите алгоритм оценки качества классификации по F1- критерию.
- 29 Запишите формулу расчета точности (precision).
- 30 Запишите формулу расчета чувствительности (recall).
- 31 Запишите формулу расчета F-критерия.
- 32 Охарактеризуйте понятие «искусственная нейронная сеть».
- 33 Дайте определение понятию «нейрон».
- 34 Охарактеризуйте эффекты обучения искусственной нейронной сети.
- 35 Изобразите кривую обучения, иллюстрирующую эффект недообученности ИНС.
- 36 Изобразите кривую обучения, иллюстрирующую эффект переобученности ИНС

Приложение 2. Оценочные средства для проведения промежуточной аттестации

а) Планируемые результаты обучения и оценочные средства для проведения промежуточной аттестации:

Код индикатора	Индикатор достижения компетенции	Оценочные средства
Код и содержание компетенции		
ПК-8.1:	Ставит задачи по адаптации или совершенствованию методов и алгоритмов для решения комплекса задач предметной области Знает: классы методов и алгоритмов машинного обучения;	<p>Примерные вопросы к зачету:</p> <ol style="list-style-type: none"> 1. Дайте определения понятию «источник данных». 2. Приведите способы классификации источников данных. 3. Охарактеризуйте понятие «открытые данные». 4. Приведите примеры источников открытых данных. 5. Что называется обучающим набором данных 6. Дайте определение «Dataset для обучения нейронной сети» 7. Перечислите способы формирования обучающих и оценочных выборок данных 8. Сформулируйте понятие регрессионного анализа. 9. Приведите и охарактеризуйте виды регрессии. 10. Приведите примеры практических задач, требующих применения регрессионного анализа. 11. Перечислите способы решения задачи регрессии. 12. Дайте определение парной регрессии. 13. Дайте определение множественной регрессии. 14. Охарактеризуйте понятие классификации данных. 15. Виды классификации данных. 16. Перечислите методы классификации данных. 17. Приведите пример практического применения классификации. 18. Приведите алгоритм применения логистической регрессии. 19. Запишите функции штрафа при логистической регрессии.
	Умеет: ставить задачи и адаптировать методы и алгоритмы машинного обучения	<p>Примерные задания для лабораторных работ</p> <ol style="list-style-type: none"> 1.

Лабораторная работа «Сбор и подготовка данных»

Общие сведения

Целями работы являются:
- ознакомление со структурой источников открытых данных, изучение способов хранения и представления данных;
- приобретение навыка построения системы сбора данных.

Задачи:

- 1 Исследование наборов данных, представленных на портале открытых данных data.gov.ru.
 - 2 Исследование наборов данных, представленных на портале data.worldbank.org.
 - 3 Построение автоматизированной системы сбора данных.
- В качестве инструментального средства используется программное обеспечение *Microsoft Excel*.

Варианты задания

Таблица 5. Задания по части 1 «Исследование портала data.gov.ru»

Вариант	Тематика	Адрес в сети Интернет
1	Государство	http://data.gov.ru/rubriki/gosudarstvo
2	Экономика	http://data.gov.ru/rubriki/ekonomika
3	Образование	http://data.gov.ru/rubriki/education
4	Здоровье	http://data.gov.ru/rubriki/zdorove
5	Экология	http://data.gov.ru/rubrics/ecology
6	Транспорт	http://data.gov.ru/rubriki/transport
7	Культура	http://data.gov.ru/rubrics/culture
8	Спорт	http://data.gov.ru/rubrics/sport
9	Строительство	http://data.gov.ru/rubriki/stroitelstvo
10	Досуг и отдых	http://data.gov.ru/rubrics/leisure-and-entertainment
11	Торговля	http://data.gov.ru/rubriki/torgovlya
12	Туризм	http://data.gov.ru/rubrics/tourism
13	Электроника	http://data.gov.ru/rubrics/electronics
14	Картография	http://data.gov.ru/rubrics/cartography
15	Безопасность	http://data.gov.ru/rubriki/bezopasnost
16	Метеоданные	http://data.gov.ru/rubrics/weather

2. Лабораторная работа «Бинарная классификация»

Общие сведения:

Целью работы является приобретение навыка бинарной классификации данных на основе логистической регрессии. В качестве инструментального средства используется программное обеспечение Microsoft Excel

Порядок выполнения:

1 Подготовка:

- 1.1 Выберите вариант задания.
- 1.2 Подготовьте выборку данных в ПО Microsoft Excel.
- 1.3 Постройте диаграмму, отображающую выборку данных.

2 Классификация:

- 2.1 Задайте целевую функцию.
- 2.2 Определите коэффициенты функции гипотезы с помощью инструмента «Поиск решения».
- 2.3 Рассчитайте значения точности, чувствительности, F-критерия.

3 Сделайте вывод об эффективности этого классификатора.

4 Отчет о работе:

- 4.1 Составьте отчет о работе.
- 4.2 Преобразуйте отчет в формат PDF.
- 4.3 Запакуйте отчет (PDF) и файл с данными (XLS) в один архив формата ZIP.

Имеет практический опыт: участия в проектах по изучению опыта

Примерные темы курсовых работ

1. Технология литейного производства
2. Специализированное оборудование для обработки

	<p>адаптации и применимости методов и алгоритмов машинного обучения для решения прикладных задач в металлургии; постановки задач по адаптации или совершенствованию методов и алгоритмов для решения комплекса задач предметной области</p>	<p>металлов давлением</p> <ol style="list-style-type: none"> 3. Оборудование и проектирование цехов и участков по получению новых материалов и изделий 4. Технология метало-обрабатывающего производства 5. Проектирование литых заготовок и оснастки 6. Проектирование цехов цветной металлургии 7. Металлургия свинца, цинка и кадмия 8. Металлургия легких металлов 9. Расчет и конструирование стальных конструкций 10. Технологии процессов обогащения полезных ископаемых согласно заданным параметрам 11. Процессы обогащения полезных ископаемых 12. Металлорежущие станки 13. Расчет теоретических балансов продуктов обогащения 14. Производство отливок из стали и чугуна 15. Свариваемость металлов и сплавов при различных методах и режимах сварки <p>Задание</p> <ol style="list-style-type: none"> 1. Найдите набор данных на одном из порталов открытых данных: data.gov.ru, data.worldbank.org, rosstat.gov.ru и др. по тематике курсовой работы. Набор должен быть представлен в формате csv кодировке Windows. Далее исходный набор данных обрабатывается в программе Excel. 2. На основе набора данных подготовьте выборку, содержащую значения выбранного показателя за все годы для трёх предприятий; постройте график, иллюстрирующий изменение показателя 3. Постройте линейную регрессию аналитическим и численными методами: <ul style="list-style-type: none"> • Выполните поиск параметров функции регрессии с помощью нормального уравнения, постройте график, создайте прогноз • Выполните поиск параметров функции регрессии с помощью инструмента «Поиск решения», постройте график, создайте прогноз • Сравните коэффициенты уравнения регрессии, полученные обоими методами, сделайте вывод 4. Выполните бинарную классификацию исходного набора методом логистической регрессии: <ul style="list-style-type: none"> • Подготовьте обучающую выборку: задайте целевую функцию, функцию штрафа, начальные значения коэффициентов.
--	---	---

		<ul style="list-style-type: none"> • Выполните численное решение с помощью инструмента «Поискрешения» • На основе численного решения сделайте предположение о принадлежности объектов одному из двух классов • Постройте таблицу, которая содержит: объекты исходной выборки, реальный класс каждого объекта, класс каждого объекта согласно выполненному алгоритму, одну из групп классификации (Truepositives TP, Falsepositives FP, Falsenegatives FN, truenegatives TN)) <p>5. Рассчитайте значения точности, чувствительности, F-критерия. Сделайте вывод об эффективности выполненной классификации</p>
--	--	--

б) Порядок проведения промежуточной аттестации, показатели и критерии оценивания:

Аттестация по дисциплине включает ответы на теоретические вопросы и выполнение практических заданий, выявляющих степень сформированности компетенций, проводится в форме зачета и зачета и курсовой работы.

Показатели и критерии оценивания зачета:

- оценка «зачтено» ставится при наборе учащимся от 3 до 5 баллов
- на оценку «не зачтено» (1-2 балла) – обучающийся не может показать знания на уровне воспроизведения и объяснения информации, не может показать интеллектуальные навыки решения простых задач