

ДНІПРОВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ОЛЕСЯ ГОНЧАРА

Факультет _____ Прикладної математики _____

Кафедра _____ Математичного забезпечення ЕОМ _____

«ЗАТВЕРДЖУЮ»

Проректор

з наукової роботи



Сергій ОКОВИТИЙ

2020 р.

РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Зу-5-15 Методи машинного навчання в інформаційних системах

для здобувачів вищої освіти

рівня вищої освіти _____ третій (освітньо-науковий) _____

галузі знань _____ 11 Математика і статистика _____

спеціальності _____ 113 Прикладна математика _____

освітньо-наукова програма _____ Прикладна математика _____

факультет _____ прикладної математики _____


вид дисципліни _____ вибіркова _____

Дніпро
2020


Розробник: Мацуга Ольга Миколаївна, канд. техн. наук, доцент, доцент кафедри математичного забезпечення ЕОМ

Робоча програма схвалена:

на засіданні кафедри аерогідромеханіки та енергомасопереносу від «10» 09 2020 року. Протокол № 2

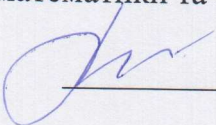
Зав. каф. аерогідромеханіки та енергомасопереносу  (Андрій ДРЕУС)

на засіданні кафедри комп'ютерних технологій від «10» 09 2020 року. Протокол № 3

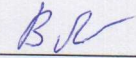
Зав. каф. комп'ютерних технологій  (Наталія ГУК)

на засіданні кафедри обчислювальної математики та математичної кібернетики від «14» 09 2020 року. Протокол № 3

Зав. каф. обчислювальної математики та математичної кібернетики

 (Валентина ТУРЧИНА)

на засіданні кафедри теоретичної та комп'ютерної механіки від «14» 09 2020 року. Протокол № 2

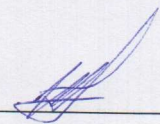
Зав. каф. теоретичної та комп'ютерної механіки  (Володимир ЛОБОДА)

Ухвалено:

на засіданні науково-методичної ради факультету прикладної математики від «14» 09 2020 року. Протокол № 3

Голова НМРФ  (Ольга ПРИТОМАНОВА)

на засіданні науково-методичної ради механіко-математичного факультету від «14» 09 2020 року. Протокол № 1

Голова НМРФ  (Олександр ГУБІН)

Робочу програму схвалено на засіданні кафедр на наступний навчальний рік

А та Е	20__/20__ н.р., протокол № _____	від «__» _____	20__ р.
КТ	20__/20__ н.р., протокол № _____	від «__» _____	20__ р.
ОМ та МК	20__/20__ н.р., протокол № _____	від «__» _____	20__ р.
ТКМ	20__/20__ н.р., протокол № _____	від «__» _____	20__ р.

1. МЕТА ДИСЦИПЛІНИ

Машинне навчання – це наука про те, як навчити машину самостійно видобувати з даних приховані закономірності та приймати рішення. Машинне навчання є невід’ємною складовою систем автоматизованої медичної діагностики, виявлення шахрайства у банках, спаму, відеофіксації правопорушень, персоналізованих рекомендацій та багатьох інших. Тому розуміння задач машинного навчання і володіння методами та засобами їх розв’язання є важливе для спеціалістів в галузі інформаційних технологій та прикладної математики.

Мета викладання дисципліни «Методи машинного навчання в інформаційних системах» – ознайомити здобувачів вищої освіти з основними задачами машинного навчання, навчити методам та засобам їх розв’язання, прищепити практичні навички їх розв’язання та надати досвід розробки програмного забезпечення для їх розв’язання на базі мови і бібліотек Python.

Дисципліна формує такі **компетентності** за освітньо-науковою програмою «Прикладна математика»:

ЗК 01. Здатність до інтелектуальної творчої діяльності, спрямованої на одержання нових знань, абстрактного мислення, аналізу та синтезу.

ЗК 02. Здатність до використання новітніх інформаційних і комунікаційних технологій, пошуку, оброблення та аналізу інформації з різних джерел.

ФК 01. Здатність до виявлення актуальних математичних проблем, використання поглиблених знань у галузі прикладної математики, внесення оригінального вкладу в її розвиток.

ФК 02. Здатність до створення адекватних математичних моделей на основі положень наукових теорій та відомостей про об’єкт дослідження.

ФК 03. Здатність обирати раціональні шляхи розв’язання поставлених задач, застосовувати сучасний математичний апарат та комп’ютерні технології.

ФК 04. Здатність застосовувати математичні моделі для дослідження складних процесів у природничих, технічних, економічних і соціальних системах.

2. ПОПЕРЕДНІ ВИМОГИ ДО ОПАНУВАННЯ АБО ВИБОРУ НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Для опанування теоретичних положень навчальної дисципліни здобувачі вищої освіти повинні мати базові знання з лінійної алгебри, теорії ймовірностей та математичної статистики, теорії оптимізації.

Для виконання практичних робіт та самостійної роботи із навчальної дисципліни здобувачі вищої освіти повинні володіти знаннями та практичними навичками розробки програмного забезпечення.

3. РЕЗУЛЬТАТИ НАВЧАННЯ ЗА ДИСЦИПЛІНОЮ ТА ЇХ СПІВВІДНОШЕННЯ ІЗ ПРОГРАМНИМИ РЕЗУЛЬТАТАМИ НАВЧАННЯ

В результаті вивчення дисципліни «Методи машинного навчання в інформаційних системах» здобувач вищої освіти повинен

знати:

- фундаментальні поняття та сфери застосування машинного навчання;
- постановки задач навчання з вчителем (класифікації, регресії) та без вчителя (кластеризації, зниження розмірності);
- методи відновлення багатовимірної регресії;

- способи вирішення проблем, що виникають під час відновлення багатовимірної регресії;
- методи розв'язання задачі класифікації, зокрема методи навчання дерев рішень, логістичної регресії та ансамблів класифікаторів;
- методи кластерного аналізу;
- підходи до оцінки якості розв'язку задач регресії, класифікації та кластеризації;
- методи та підходи до зниження розмірності даних;

вміти:

- формулювати постановку задачі машинного навчання, у тому числі в рамках власного наукового дослідження;
- обґрунтовано обирати та застосовувати на практиці методи машинного навчання;
- робити висновки щодо якості розв'язання задач машинного навчання та надавати інтерпретацію отриманим результатам;
- розробляти програмне забезпечення для розв'язання задач машинного навчання.

Наведені результати навчання за дисципліною співвідносяться із такими **програмними результатами навчання:**

ПР 01. Знати на поглибленому рівні фундаментальні моделі, методи та алгоритми прикладної математики.

ПР 03. Уміти обґрунтовувати вибір математичної моделі на основі інтелектуального аналізу даних про об'єкт дослідження та наявного спектру моделей.

ПР 04. Уміти розробляти методики та обчислювальні алгоритми математичного та комп'ютерного моделювання складних природничих, технічних, економічних і соціальних систем.

ПР 05. Уміти розробляти програмне забезпечення для реалізації алгоритмів моделювання складних систем і процесів.

4. СТРУКТУРА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Форма навчання денна

2 семестр

№ п/п	Номер і назва теми	Кількість годин				Примітки			
		Лекції	Практичні	Лабораторні заняття	Самостійна робота	2020/21 н.р.	2021/22 н.р.	2022/23 н.р.	2023/24 н.р.
Розділ 1. Методи машинного навчання в інформаційних системах									
1	Тема 1. Вступ до машинного навчання	2			6				
2	Тема 2. Задача відновлення моделі багатовимірної регресії. Проблеми під час відновлення багатовимірної регресії. Оцінка якості відновленої регресії	4	4		18				
3	Тема 3. Задача класифікації. Оцінка якості класифікації	2	2		8				
4	Тема 4. Дерева рішень	2	2		14				
5	Тема 5. Логістична регресія	2	2		10				
6	Тема 6. Ансамблі класифікаторів	2	2		14				
7	Тема 7. Задача кластеризації. Оцінка якості кластеризації	2	2		10				
8	Тема 8. Методи кластеризації	2	2		14				
9	Тема 9. Зниження розмірності. Методи відбору ознак. Метод головних компонент	2	2		18				
	ВСЬОГО	20	18		112				

5. СХЕМА ФОРМУВАННЯ ОЦІНКИ

5.1 Шкала відповідності оцінювання

Відмінно/Excellent	Зараховано/Passed	90-100
Добре/Good		82-89
Задовільно/Satisfactory		75-81
		64-74
Незадовільно/Fail	Не зараховано/Fail	60-63
		0-59

5.2 Форми та організація оцінювання

Поточне оцінювання:

Форма оцінювання	Терміни оцінювання (тиждень)	Максимальна кількість балів
Виконання практичної роботи на тему «Відновлення регресійної моделі»	25	15
Виконання практичної роботи на тему «Розв'язання задачі класифікації»	31	15
Виконання практичної роботи на тему «Розв'язання задачі кластеризації»	37	15
Виконання самостійної роботи на тему «Методи машинного навчання у власному дисертаційному дослідженні»: звіт та захист.	37	55
Максимальна кількість балів за поточне оцінювання		100

Підсумкове оцінювання:

Форма оцінювання	Терміни оцінювання (тиждень)	Максимальна кількість балів
Диференційований залік	38	100

6. ІНСТРУМЕНТИ, ОБЛАДНАННЯ ТА ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ, ВИКОРИСТАННЯ ЯКИХ ПЕРЕДБАЧАЄ НАВЧАЛЬНА ДИСЦИПЛІНА

Навчальна дисципліна передбачає використання програмно-апаратних та інструментальних засобів розробки програмного забезпечення. В якості мови програмування передбачається використання Python, допускається також використання мов C, C++, C#, Java, R тощо.

7. РЕКОМЕНДОВАНА ЛІТЕРАТУРА

Основна (Базова):

1. Bishop C.M. Pattern Recognition and Machine Learning / C.M. Bishop. – Springer, 2006. – 738 p.
2. Hastie T. The Elements of Statistical Learning. Data Mining, Inference, and Prediction / T. Hastie, R. Tibshirani, J. Friedman. – 2009. – 745 p.
3. James J. An Introduction to Statistical Learning with Applications in R / G. James, D. Witten, T. Hastie, R. Tibshirani. – Springer, 2013. – 426 p.
4. Murphy K.P. Machine learning. A probabilistic perspective / K.P. Murphy. – Massachusetts Institute of Technology, 2012. – 1098 p.
5. Бабак, В.П. Статистична обробка даних. / В.П. Бабак, А.Я. Білецький, О.П. Приставка, П.О. Приставка. – К.: МІВВІЦ, 2001. – 388 с.
6. Мацуга О.М. Навчальний посібник до вивчення курсу «Інформаційні технології розпізнавання образів» / О.М. Мацуга, Ю.М. Архангельська, Н.М. Єрещенко. – Д.: РВВ ДНУ, 2016. – 60 с.

Допоміжна:

1. Воронцов К.В. Машинное обучение (курс лекций) [Электронный ресурс] / К.В. Воронцов. – Режим доступа: [http://www.machinelearning.ru/wiki/index.php?title=Машинное_обучение_\(курс_лекций%2С_К.В.Воронцов\)](http://www.machinelearning.ru/wiki/index.php?title=Машинное_обучение_(курс_лекций%2С_К.В.Воронцов)).
2. Дрейпер Н. Прикладной регрессионный анализ / Н. Дрейпер, Г. Смит. – М.: Вильямс, 2007. – 912 с.
3. Coelho L.P. Building Machine Learning Systems with Python. 2nd ed. / L.P. Coelho, W. Richert. – Packt Publishing Ltd, 2015. – 301 p.
4. Harrington P. Machine Learning in Action / P. Harrington. – Manning, 2012. – 354 p.
5. Jain A.K. Data clustering: A review / A.K. Jain, M.N. Murty, P.J. Flynn // ACM Computing Surveys. – 1999. – Vol. 31, № 3. – P. 264–323. – Режим доступа <https://www.cs.rutgers.edu/~mlittman/courses/lightai03/jain99data.pdf>
6. Marsland S. Machine Learning. An Algorithmic Perspective. 2nd ed. / S. Marsland. – CRC Press, 2015. – 430 p.
7. Zaki M.J. Data Mining and Analysis. Fundamental Concepts and Algorithms / M.J. Zaki, Meira Jr. W. – Cambridge University Press, 2014. – 593 p.
8. Айвазян С.А. Классификация многомерных наблюдений / С.А. Айвазян, З.И. Бежаева, О.В. Староверов. – М.: Статистика, 1974. – 240 с.
9. Айвазян С.А. Прикладная статистика: исследование зависимости / С.А. Айвазян, И.С. Енюков, Л.Д. Мешалкин. – М.: Финансы и статистика, 1983. – 471 с.
10. Бериков В.С. Современные тенденции в кластерном анализе / В.С. Бериков, Г.С. Лбов. // Всероссийский конкурсный отбор обзорно-аналитических статей по приоритетному направлению «Информационно-телекоммуникационные системы», 2008. – 26 с.
11. Горелик А.Л. Методы распознавания / А.Л. Горелик, В.А. Скрипник. – М.: Высш. шк., 1984. – 208 с.
12. Загоруйко Н.Г. Прикладные методы анализа данных и знаний. / Н.Г. Загоруйко. – Новосибирск: ИМ СО РАН, 1999. – 270 с.
13. Ким Дж.-О. Факторный, дискриминантный и кластерный анализ / Дж.-О. Ким, Ч.У. Мьюллер, У.Р. Клекка, М.С. Олдендерфер, Р.К. Блэшфилд. – М.: Финансы и статистика, 1989. – 215 с.
14. Компьютерная биометрика / Под ред. В.Н. Носова. – М.: Изд-во МГУ, 1990. – 232 с.
15. Мандель И.Д. Кластерный анализ / И.Д. Мандель. – М.: Финансы и статистика, 1988. – 176 с.
16. Паклин Н.Б. Бизнес-аналитика: от данных к знаниям (+CD): Учеб. пособ. 2-е изд., испр. / Н.Б. Паклин, В.И. Орешков. – СПб: Питер, 2013. – 704 с.
17. Приставка П.О. Анализ даних / П.О. Приставка, О.М. Мацуга. – Д.: РВВ ДНУ, 2008. – 92 с.
18. Шахиди А. Деревья решений [Электронный ресурс] / А. Шахиди. – Режим доступа: <http://www.basegroup.ru/library/analysis/tree/>
19. Фукунага К. Введение в статистическую теорию распознавания образов / К. Фукунага. – М.: Наука, 1979. – 368 с.

8. ІНФОРМАЦІЙНІ РЕСУРСИ

1. <http://repository.dnu.dp.ua>
2. <https://habr.com/ru/company/ods/blog/322626/>
3. <https://scikit-learn.org/stable/index.html>
4. <https://archive.ics.uci.edu/ml/datasets.html>

5. <https://www.kaggle.com/>