

МІ МА

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Харківський національний автомобільно-дорожній університет

«ЗАТВЕРДЖУЮ»

Заступник ректора ХНАДУ

професор  Гладкий І.П.

« » 2017 року



exam

ПРОГРАМА

навчальної дисципліни	<u>Інтелектуальний аналіз даних</u> (назва навчальної дисципліни згідно навчального плану)
підготовки	<u>бакалавра</u> (назва освітньо-кваліфікаційного рівня)
галузі знань	<u>12 Інформаційні технології</u> (шифр і назва галузі знань)
спеціальності	<u>121 Інженерія програмного забезпечення</u> (шифр і назва напрямку підготовки)
кваліфікація	<u>3121 Фахівець з інформаційних технологій</u> (шифр і назва кваліфікації для бакалавра, спеціальності – для магістра)

(шифр № 1.3.05)

2017 рік

Розроблено та внесено: кафедрою комп'ютерних технологій та мехатроніки

Розробники програми: професор кафедри комп'ютерних технологій та мехатроніки, д.т.н., доцент Клець Дмитро Михайлович

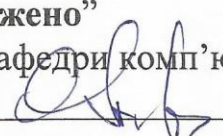
Обговорено та рекомендовано до затвердження на засіданні кафедри.
Протокол № 1 від 28.08.2017 р.

Завідувач кафедри докт. техн. наук, доцент



Клець Д.М.

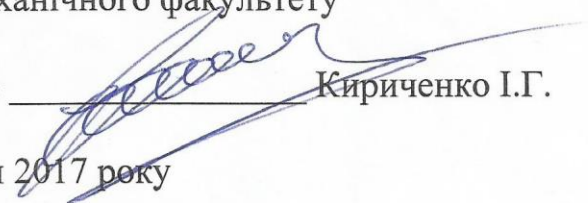
✓ **“Узгоджено”**

Завідувач кафедри комп'ютерних технологій та мехатроніки
доцент  Клець Д.М.

28 серпня 2017 року

“Узгоджено”

Декан механічного факультету

професор 

Кириченко І.Г.

28 серпня 2017 року

ВСТУП

Програма вивчення навчальної дисципліни “Інтелектуальний аналіз даних” складена відповідно до освітньо-кваліфікаційної характеристики та навчального плану підготовки **бакалавра спеціальності 121 Інженерія програмного забезпечення**.

1. Мета, предмет та завдання навчальної дисципліни

1.1. **Метою** викладання дисципліни є теоретична та практична підготовка студентів до вивчення систем обробки даних та принципів інтелектуального аналізу даних на основі методів та алгоритмів Data Mining.

1.2. **Предмет дисципліни** включає створення та дослідження математичних та програмних моделей інтелектуального аналізу даних, пов'язаних із функціонуванням об'єктів професійної діяльності.

1.3. **Основними завданнями** є формування у студентів уяви про застосування технологій зберігання та організації даних, методів та алгоритмів Data Mining, процесів виявлення знань, принципів побудови сховищ даних, а також способів візуального представлення даних.

1.4. Очікувані результати навчання з дисципліни:

По завершенні вивчення дисципліни студенти повинні володіти наступними загальними компетентностями:

- здатність до абстрактного мислення, аналізу та синтезу;
- здатність застосовувати знання у практичних ситуаціях;
- здатність спілкуватися державною мовою як усно, так і письмово;
- здатність до пошуку, оброблення та аналізу інформації з різних джерел.

Та наступними фаховими компетентностями:

- здатність створювати програмне забезпечення для зберігання, видобування та опрацювання даних;
- здатність накопичувати, обробляти та систематизувати професійні знання щодо створення і супроводження програмного забезпечення та визнання важливості навчання протягом всього життя.

Результати навчання:

- розуміти, аналізувати, цілеспрямовано шукати і вибирати необхідні для вирішення професійних завдань інформаційно-довідникові ресурси і знання з урахуванням сучасних досягнень науки і техніки;
- знати та вміти використовувати методи та засоби збору, формулювання та аналізу вимог до програмного забезпечення;
- знати, розуміти і застосовувати ефективні підходи щодо проектування програмного забезпечення;
- знати, розуміти і застосовувати професійні стандарти і інші нормативно-правові документи в галузі інженерії програмного забезпечення;
- знати, розуміти і застосовувати відповідні математичні поняття, методи доменного, системного і об'єктно-орієнтованого аналізів та математичного моделювання для розробки програмного забезпечення.

Міждисциплінарні зв'язки:

Вивченню дисципліни

передують: «Вища математика», «Дискретна математика», «Теорія ймовірностей, ймовірні процеси», «Теорія алгоритмів», «Чисельні методи», «Математичні

методи дослідження операцій», «Організація баз даних та знань». *потребують її вивчення: «Проектування інформаційних систем», «Методи та системи штучного інтелекту», «Системна інженерія», «Моделювання транспортних потоків», «Переддипломна практика», «Дипломне проектування».*

1. ОПИС НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Найменування показників	Характеристика навчальної дисципліни	
	денна форма навчання	заочна (дистанційна) форма навчання ¹
Кількість кредитів – 4 Кількість годин - 120	<u>нормативна</u> (нормативна, за вибором ВНЗ, за вибором студента)	
Семестр викладання дисципліни	7	
Вид контролю:	<u>екзамен</u> (залік, екзамен)	
Розподіл часу:		
- лекції (годин)	16	
- практичні, семінарські (годин)		
- лабораторні роботи (годин)	32	
- самостійна робота студентів (годин)	72	
- курсовий проект (годин)		
- курсова робота (годин)		
- розрахунково-графічна робота (контрольна робота)		
- підготовка та складання екзамену		

2. ІНФОРМАЦІЙНИЙ ОБСЯГ НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Розділ 1. Первинна обробка даних

Тема 1. Основи інтелектуального аналізу даних.

Поняття інтелектуального аналізу даних. Методи, стадії та завдання інтелектуального аналізу даних. Data Mining як частина ринку інформаційних технологій. Розвиток технології баз даних. Порівняння статистики, машинного навчання та Data Mining.

Тема 2. Регресійний аналіз та нелінійні моделі

Методи первинної обробки даних. Методи дослідження структури даних: візуалізація та автоматичне групування даних. Описова статистика. Кореляційний аналіз. Регресійний аналіз. Множинний регресійний аналіз. Нелінійна множинна регресійна модель. Нелінійне оцінювання параметрів.

Тема 3. Методи багатомірного розвідувального аналізу

Методи кластерного аналізу. Заходи подібності. Методи об'єднання або зв'я-

¹ Якщо дисципліна на заочній (дистанційній) формі навчання не викладається, то графа “заочна форма навчання” відсутня.

зку. Ієрархічний кластерний аналіз в SPSS. Ітеративні методи. Алгоритм k-середніх (k-means). Алгоритм РАМ. Попереднє скорочення розмірності. Факторний аналіз. Ітеративна кластеризація в SPSS. Растрова кластеризація об'єктів. Лінійний дискримінантний аналіз.

Розділ 2. Класифікація та прогнозування

Тема 4. Методи класифікації

Завдання класифікації. Процес класифікації. Методи, що застосовуються для вирішення задач класифікації. Точність класифікації: оцінка рівня помилок. Оцінювання класифікаційних методів. Дерева рішень. Методи опорних векторів, «найближчого сусіда», або системи міркувань на основі аналогічних випадків. Метод Байєса. Лінійний SVM. Класифікація об'єктів у випадку невідомих розподілень даних.

Тема 5. Методи прогнозування

Завдання прогнозування. Порівняння задач прогнозування та класифікації. Прогнозування і тимчасові ряди. Завдання та методи візуалізації. Рішення завдання прогнозування. Аналіз багатомірних угруповань. Статистична обробка тимчасових рядів і прогнозування.

Тема 6. Методи пошуку шаблонів даних

Асоціативні правила. Послідовне відображення шаблонів даних. Методи пошуку асоціативних правил. Додатки із застосуванням асоціативних правил. Характеристики асоціативних правил. Межі підтримки і достовірності асоціативного правила. Методи пошуку асоціативних правил. Різновиди алгоритму Apriori.

Розділ 3. Упровадження моделі та представлення даних

Тема 7. OLAP та Data Mining

Методи, стадії задачі Data Mining. Упровадження Data Mining, OLAP і сховищ даних у СППР. Процес Data Mining. Стандарти Data Mining. Інструменти Data Mining.

Тема 8. Способи візуального представлення даних

Візуалізація інструментів Data Mining. Візуалізація Data Mining моделей. Методи візуалізації. Представлення даних в одному, двох і трьох вимірах. Паралельні координати. Якість візуалізації. Подання просторових характеристик. Основні тенденції в області візуалізації.

Теми лабораторних занять:

1. Первинний розвідувальний аналіз даних (4 години).
2. Кластерний аналіз даних (4 години).
3. Застосування дерев класифікації у вирішенні завдань інтелектуального аналізу даних (4 години).

4. Прогнозування часових рядів (4 години).
5. Кореляційно-регресивний аналіз (4 години).
6. Обробка та аналіз текстових даних за допомогою Text Miner (4 години).
7. Описова і передбачувальна аналітика за допомогою платформи Deductor Studio Academic (4 години).
8. Звітність та візуалізація оброблених даних за допомогою платформи Deductor Studio Academic (4 години).

3. РЕКОМЕНДОВАНА ЛІТЕРАТУРА ТА ІНФОРМАЦІЙНІ РЕСУРСИ

Базова

1. Барсегян А.А. Анализ данных и процессов / Барсегян А.А., Куприянов М.С., Холод И.И., Тесс М.Д., Елизаров С.И. // 3-е изд. – СПб.: БХВ-Петербург, 2009. – 512 с. – ISBN: 978-5-9775-0368-6
2. Чубукова И. А. Data Mining: учебное пособие / И.А. Чубукова. - 2-е изд., испр. – М.: INTUIT; БИНОМ. Лаборатория знаний, 2008. – 382 с. – ISBN: 978-5-9556-0064-2.
3. Дюк В. Data Mining: учебный курс / Дюк В., Самойленко А. – СПб: Изд. Питер 2001. — 368 с. – ISBN: 5-318-00227-7.
4. Барсегян А.А. Технологии анализа данных. Data Mining, Visual Mining, Text Mining, OLAP / Барсегян А.А., Куприянов М.С., Степаненко В.В., Холод И. И. // 2-е изд. – СПб.: БВХ–Петербург, 2007, – 384 с. – 3-е изд. – ISBN 5-94157-991-8.
5. Барсегян А. А. Методы и модели анализа данных OLAP и Data Mining / Барсегян А. А., Куприянов М. С., Степаненко В. В., Холод И. И. // СПб.: БВХ–Петербург, 2004, – 336 с. – ISBN 5-94157-522-X.

Допоміжна

6. Спирли Эрик – Корпоративные хранилища данных. Планирование, разработка и реализация. Том 1 – Пер с англ. – М.: Издательский дом «Вильямс», 2001. – 400с.: ил.
7. Корнеев В. В. Базы данных. Интеллектуальная обработка информации / В. В. Корнеев, А. Ф. Гареев, С. В. Васютин и др. – М. : Издатель Могачева С. В. ; Издательство Нолидж, 2001. – 496 с.
8. Рассел С. Искусственный интеллект: современный подход / Рассел С., Норвинг П. – М. : Издательский дом "Вильямс", 2006. – 1408 с.

4. ЗАСОБИ ДІАГНОСТИКИ УСПІШНОСТІ НАВЧАННЯ

(вказати перелік засобів контролю успішності навчання студентів, які застосовуються: тести, екзаменаційні білети, тощо)

Підсумковий контроль знань та компетентностей студентів з навчальної дисципліни здійснюється на підставі захисту курсової роботи та проведення семестрового екзамену.

Екзаменаційний білет включає теоретичні питання та практичні завдання, які оцінюються відповідно до положення "Про порядок оцінювання результатів навчання студентів за накопичувальною бально-рейтинговою системою"

ХНАДУ.

До основних завдань контролю знань студентів в університеті належать:

- оцінювання рівня засвоєння студентами програм навчальних дисциплін та інформування студентів про якість досягнених результатів;
- мотивація студентів до систематичної активної роботи протягом усього періоду навчання;
- аналіз успішності та вплив викладача на процес самостійної роботи студентів і ефективність навчального процесу в цілому.

Дидактичними принципами системи контролю знань студентів є: дієвість; систематичність; індивідуальність; диференційованість; об'єктивність; єдність вимог; прозорість навчального середовища.

Зазначені принципи контролю логічно пов'язані між собою і визначають вимоги до форм і методів перевірки та оцінювання знань, що формують систему контролю знань студентів.

Реалізація основних завдань контролю знань студентів у ХНАДУ досягається системними підходами до оцінювання та комплексністю застосування різних видів контролю. За місцем, яке посідає контроль у навчальному процесі, розрізняють: вхідний контроль, поточний контроль, модульний контроль, семестровий підсумковий контроль (залік або екзамен), державну атестацію та контроль залишкових знань (ректорський контроль).

Критеріями оцінки знань є засвоєння теоретичних основ та розуміння практичних аспектів; обсяг знань та ступінь розуміння матеріалу; самостійність мислення; знання законодавчої бази з означених питань; логічність мислення та активність в процесі проведення занять. Критерії оцінки знань студентів за шкалою оцінювання ХНАДУ наведені в табл. 1.

Таблиця 1 - Критерії оцінки знань студентів

Кількість балів	Критерії
90 - 100	Студент володіє узагальненими знаннями навчального матеріалу в повному обсязі та здатний їх ефективно використовувати для виконання всіх передбачених навчальною програмою практичних завдань. Відповідь студента повна, правильна, логічна і містить аналіз, систематизацію, узагальнення навчального матеріалу. Вміє самостійно знаходити і користуватися джерелами інформації, оцінювати отриману інформацію. Встановлює причинно-наслідкові та між предметні зв'язки. Робить аргументовані висновки. Правильно і усвідомлено застосовує всі види додаткової інформації. Практичні завдання виконує правильно у повному обсязі. Виказує пізнавально-творчий інтерес до предмету.
80 - 89	Недостатньо повно та ґрунтовно засвоїв окремі питання робочої програми. Вміє самостійно викласти зміст основних питань програми навчальної дисципліни, виконав завдання кожної теми та модульного поточного контролю в цілому.
75 - 79	Недостатньо повно та ґрунтовно засвоїв деякі теми робочої програми, не вміє самостійно викласти зміст деяких питань програми навчальної дисципліни. Окремі завдання кожної теми та модульного поточного контролю в цілому виконав не повністю.
67 - 74	Засвоїв лише окремі теми робочої програми. Не вміє вільно самостійно викласти зміст основних питань навчальної дисципліни, окремі завдання кожної теми модульного контролю не виконав.

Кількість балів	Критерії
60 - 66	Засвоїв лише окремі питання навчальної програми. Не вміє достатньо самостійно викласти зміст більшості питань програми навчальної дисципліни. Виконав лише окремі завдання кожної теми та модульного контролю в цілому.
35 - 59	Не засвоїв більшості тем навчальної програми не вміє викласти зміст більшості основних питань навчальної дисципліни. Не виконав більшості завдань кожної теми та модульного контролю в цілому.
1 - 34	Не засвоїв навчальної програми, не вміє викласти зміст кожної теми навчальної дисципліни, не виконав модульного контролю.

Отримана студентом загальна оцінка може бути переведена з рейтингової системи ВНЗ в інші системи оцінки знань студента за наступною системою (табл. 2).

Таблиця 2 - Критерії оцінки знань студентів за національною шкалою оцінювання

Кількість балів	Оцінка за шкалою ECTS	Оцінка за національною шкалою	
		для іспиту, курсової роботи, практики	для заліку
90 - 100	A – відмінно	5 – відмінно	зараховано
80 - 89	B – дуже добре	4 – добре	
75 - 79	C – добре		
67 - 74	D – задовільно	3 – задовільно	
60 - 66	E – достатньо		
35 - 59	FX – незадовільно з можливістю повторного складання	2 – незадовільно	не зараховано
1 - 34	F – незадовільно з обов’язковим повторним вивченням дисципліни		

У разі отримання незадовільної оцінки, перескладання екзамену (заліку) з дисципліни допускається не більше двох разів. Перший раз – науково-педагогічному працівнику, при другому перескладанні – в усній чи комбінованій формі комісії, яка створюється деканом. Оцінка комісії є остаточною.

Розробник програми:
професор кафедри комп’ютерних технологій
та мехатроніки, докт. техн. наук, доцент

Клец Д.М.