Форма №4

ЗАПОРІЗЬКИЙ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ
(повне найменування вищого навчального закладу)
Кафедра Програмних засобів

“**ЗАТВЕРДЖУЮ**”
Ректор (перший проректор)

_____В.Г. Прушківський
“_____”_____2016 року

РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

ІНТЕЛЕКТУАЛЬНИЙ АНАЛІЗ ДАНИХ
(код і назва навчальної дисципліни)
напрям підготовки 6.050103 «Програмна інженерія»
(код і назва напряму підготовки)
спеціальність «Програмне забезпечення систем»
(код і назва спеціальності)
спеціалізація _____
(назва спеціалізації)
інститут, факультет Інститут інформатики та радіоелектроніки, факультет комп’ютерних наук і технологій
(назва інституту, факультету)

Запоріжжя – 2016 рік

Робоча програма ___«Інтелектуальний аналіз даних»___ для студентів
(назва навчальної дисципліни)
за напрямом підготовки 6.050103 «Програмна інженерія», спеціальністю «Програмне забезпечення систем». „___” _____, 2016 року - 9 с.

_____

Розробники: (вказати авторів, їхні посади, наукові ступені та вчені звання)
Федорончак Тетяна Василівна, доц. каф. ПЗ, к.т.н., доц.

Робоча програма затверджена на засіданні кафедри
програмних засобів

Протокол від. “____”_____2016 року № ___

Завідувач кафедри програмних засобів

_____ (___Субботін С.О.___)
(підпис) (прізвище та ініціали)
“_____”_____ 2016 року

Схвалено методичною комісією вищого навчального закладу за напрямом підготовки (спеціальністю) 6.050103 «Програмна інженерія»
(код, назва)

Протокол від. “____”_____2016 року № ___

“_____”_____20__ року Голова _____ (___М.М. Касьян___)
(підпис) (прізвище та ініціали)

©Федорончак Т.В., 2016 рік
© ЗНТУ, 2016 рік

1. Опис навчальної дисципліни

<table>
<tr><td rowspan="2">Найменування показників</td><td rowspan="2">Галузь знань, напрям підготовки, освітньо-кваліфікаційний рівень</td><td colspan="2">Характеристика навчальної дисципліни</td></tr>
<tr><td>денна форма навчання</td><td>заочна форма навчання</td></tr>
<tr><td rowspan="2">Кількість кредитів – 2</td><td>Галузь знань
0501 «Інформатика та обчислювальна техніка»
(шифр і назва)</td><td colspan="2" rowspan="2">За вибором</td></tr>
<tr><td>Напрям підготовки
6.050103 «Програмна інженерія»
(шифр і назва)</td></tr>
<tr><td>Модулів – 2</td><td rowspan="5">Спеціальність (професійне спрямування):
Програмне забезпечення систем</td><td colspan="2">Рік підготовки:</td></tr>
<tr><td>Змістових модулів – 5</td><td>4-й</td><td>4-й</td></tr>
<tr><td>Індивідуальне науково-дослідне завдання
_____
(назва)</td><td colspan="2">Семестр</td></tr>
<tr><td rowspan="2">Загальна кількість годин – 60</td><td>7-й</td><td>7-й</td></tr>
<tr><td colspan="2">Лекції</td></tr>
<tr><td rowspan="8">Тижневих годин для денної форми навчання:
аудиторних – 2
самостійної роботи студента – 2,3</td><td rowspan="8">Освітньо-кваліфікаційний рівень: бакалавр</td><td>14 год.</td><td>4 год.</td></tr>
<tr><td colspan="2">Практичні, семінарські</td></tr>
<tr><td>--</td><td>--</td></tr>
<tr><td colspan="2">Лабораторні</td></tr>
<tr><td>14 год.</td><td>2 год.</td></tr>
<tr><td colspan="2">Самостійна робота</td></tr>
<tr><td>32 год.</td><td>54 год.</td></tr>
<tr><td colspan="2">Індивідуальні завдання:
--</td></tr>
<tr><td colspan="2">Вид контролю: залік</td></tr>
</table>

Примітка.
Співвідношення кількості годин аудиторних занять до самостійної і індивідуальної роботи становить:

для денної форми навчання – 0,88
для заочної форми навчання – 0,11

2. Мета та завдання навчальної дисципліни

Мета – вивчення методів сучасної обробки даних – інтелектуального аналізу даних (Data Mining, Knowledge Discovery in Data), аналітичного дослідження великих масивів інформації з метою виявлення нових раніше невідомих, практично корисних знань і закономірностей, необхідних для прийняття рішень; огляд методів, програмних продуктів і різних інструментальних засобів, які використовуються Data Mining; розгляд практичних прикладів застосування Data Mining; підготовка студентів до самостійної роботи з вирішення задач засобами Data Mining і розробки інтелектуальних систем.

Data Mining – мультидисциплінарна область, яка виникла і розвивається на базі таких наук як прикладна статистика, розпізнавання образів, штучний інтелект, теорія баз даних та ін.

Завдання дисципліни полягає в тому, що студенти повинні :

− опанувати базові принципи побудови моделей даних;

− ознайомитися з концепцією Knowledge Discovery in Data (виявлення знань в даних) и Data Mining («видобування» знань);

− навчитися ефективно використовувати методи здобуття знать з великих масивів даних;

− ознайомитися з основними типами задач, що можуть бути вирішені за допомогою методів інтелектуального аналізу даних;

− отримати практичні навички з використання інструментальних засобів інтелектуального аналізу даних при вирішенні прикладних задач та навчитися інтерпретувати отримані результати.

У результаті вивчення навчальної дисципліни студент повинен

з н а т и :

− основні поняття, задачі та стадії інтелектуального аналізу даних;

− підходи к збереженню, представленню та обробці інформації в сучасних інформаційних системах;

− методи побудови моделей та аналізу залежностей у великих масивах даних;

− сучасні програмні засоби для проектування і розробки систем інтелектуального аналізу даних;

в м і т и :

− обґрунтовувати вибір конкретного типу моделі та методу інтелектуального аналізу даних при вирішенні поставленої практичної задачі;

− проводити необхідну попередню обробку даних, визначати тип задачі аналізу, вирішувати її адекватно обраним методом з оптимально визначеними параметрами, оцінювати результати, робити змістовні висновки та інтерпретацію;

− використовувати сучасні програмні засоби для проектування та дослідження систем інтелектуального аналізу даних.

3. Програма навчальної дисципліни

Змістовий модуль 1. Основи інтелектуального аналізу даних: процес виявлення знань

Тема 1. Основи інтелектуального аналізу даних**.**

Визначення Data Mining і область застосування. Задачі, моделі та методи Data Mining. Методи, стадії, задачі Data Mining. Поняття Business Intelligence.

Тема 2. Процес виявлення знань.

Цикл одержання, попередньої обробки, аналізу даних, інтерпретації результатів та їхнього використання. Етапи процесу Data Mining, пов'язані з побудовою, перевіркою, оцінкою, вибором и корекцією моделей. Методи первісної обробки даних. Інструментальні засоби Data Mining. Методи дослідження структури даних: візуалізація даних.

Змістовий модуль 2. Алгоритми Data Mining: класифікація і регресія

Тема 1. Вирішення задачі класифікації

Постановка задачі класифікації та представлення результатів. Методи побудови правил класифікації. Методи побудови дерев рішень. Методи побудови математичних функцій. Методи опорних векторів, «найближчого сусіда», Байеса. Аналіз багатомірних угрупувань. Класифікація об'єктів у випадку невідомих розподілень даних. Методи оцінювання помилок класифікації.

Тема 2. Вирішення задачі регресії

Сутність задачі прогнозування. Методи вирішення задачі регресії.

Змістовий модуль 3. Інтелектуальний аналіз часових рядів

Тема 1. Методи аналізу часових рядів

Поняття нечітких часових рядів. Методи моделювання часових рядів. Методи аналізу та прогнозування поведінки часових рядів.

Змістовий модуль 4. Алгоритми Data Mining: кластеризація

Тема 1. Вирішення задачі кластеризації

Постановка задачі кластеризації та представлення результатів. Види кластерів. Міри близькості, засновані на відстанях. Базові алгоритми кластеризации. Адаптивні методи кластеризации.

Змістовий модуль 5. Алгоритми Data Mining: пошук асоціативних правил

Тема 1. Вирішення задачі пошуку асоціативних правил

Постановка задачі пошуку асоціативних правил та представлення результатів. Секвенціальний аналіз. Різновиди задач пошуку асоціативних правил. Методи подання результатів. Алгоритми пошуку асоціативних правил. Методи пошуку асоціативних правил: метод Apriori, побудова FP-дерев пошуку шаблонів даних. Побудова hash-дерев.

4. Структура навчальної дисципліни

Назви змістових модулів і тем	Кількість годин											
	денна форма						заочна форма					
	усього	у тому числі					усього	у тому числі				
		л	п	лаб	інд	с.р.		л	п	лаб	інд	с.р.
Модуль 1												
Змістовий модуль 1. Основи інтелектуального аналізу даних, процес виявлення знань												
Тема 1. Основи інтелектуального аналізу даних	2	1	-	-	-	1	2	0,5	-	-		1,5
Тема 2. Процес виявлення знань	3	1	-	2	-	0	3	0,5	-	2		0,5
Разом за змістовим модулем 1	5	2	-	2	-	1	5	1	-	2		2
Змістовий модуль 2. Алгоритми Data Mining: класифікація і регресія												
Тема 1. Вирішення задачі класифікації	15	4	-	4	-	7	15	2	-	-		13
Тема 2. Вирішення задачі регресії	5	1		2		2	5		-	-		5
Разом за змістовим модулем 2	20	5	-	6	-	9	20	2	-	-		18
Змістовий модуль 3. Інтелектуальний аналіз часових рядів												
Тема 1. Методи аналізу часових рядів	10	1	-	-	-	9	10	-	-	-		10
Разом за змістовим модулем 3	10	1	-	-	-	9	10	-	-	-		10
Усього годин	35	8	-	8	-	19	35	3	2	-		30
Модуль 2												
Змістовий модуль 4. Алгоритми Data Mining: кластеризація												
Тема 1. Вирішення задачі кластеризації	15	4	-	4		7	15	1	-	-		14
Разом за змістовим модулем 4	15	4	-	4	-	7	15	1	-	-		14
Змістовий модуль 5. Алгоритми Data Mining: пошук асоціативних правил												
Тема 1. Вирішення задачі пошуку асоціативних правил	10	2	-	2	-	6	10		-	-		10
Разом за змістовим модулем 5	10	2	-	2		6	10		-	-		10
Усього годин	25	6	-	6	-	13	25	1	-	-		24
Разом	60	14	-	14	-	32	60	4	-	2		54

5. Теми семінарських занять

№ з/п	Назва теми	Кількість годин
	Не передбачено	

6. Теми практичних занять

№ з/п	Назва теми	Кількість годин
	Не передбачено	

7. Теми лабораторних занять

№ з/п	Назва теми	К-ть годин
1	Знайомство з програмою інтелектуального аналізу даних WEKA та підготовка даних.	2
2	Задача класифікації	4
3	Прогнозування, задача регресії	2
4	Задача кластеризації	4
5	Пошук асоціативних правил	2
	Усього	14

8. Самостійна робота

Самостійна робота студента:

(денна форма навчання) повторення матеріалу, засвоєного на лекціях, самостійне опанування частини теоретичного матеріалу, робота з контрольними запитаннями та завданнями, підготовка до лабораторних робіт, підготовка до заліку – 32 год.

(заочна форма навчання) повторення матеріалу, засвоєного на лекціях, самостійне опанування частини теоретичного матеріалу, робота з контрольними запитаннями та завданнями – 54 год.

На самостійну роботу студентів виносяться наступні теми:

№ з/п	Назва теми	К-ть годин
1	Сховища даних	5
2	Оперативний аналіз даних. OLAP-куби	5
	Разом	10

9. Індивідуальні завдання

Не передбачено.

10. Методи навчання

Організаційні форми навчання: лекційні заняття, лабораторні заняття, самостійна робота студентів.

Основні методи активного навчання: розповідь, бесіда, дискусія, діалог, індивідуальний/груповий проект, презентація.

11. Методи контролю

Усний контроль у вигляді індивідуального та фронтального опитування.

Письмовий контроль у вигляді самостійних письмових робіт, поточного тестування.

12. Розподіл балів, які отримують студенти

Поточне тестування та самостійна робота							Сума
ЗМ 1		ЗМ 2		ЗМ 3	ЗМ 4	ЗМ 5	
Т1	Т2	Т1	Т2	Т1	Т1	Т1	
10	10	20	10	15	15	20	100

Шкала оцінювання: національна та ECTS

Сума балів за всі види навчальної діяльності	Оцінка ECTS	Оцінка за національною шкалою	
		для екзамену, курсового проекту (роботи), практики	для заліку
90 – 100	**A**	відмінно	зараховано
85-89	**B**	добре	
75-84	**C**		
70-74	**D**	задовільно	
60-69	**E**		
35-59	**FX**	незадовільно з можливістю повторного складання	не зараховано з можливістю повторного складання
1-34	**F**	незадовільно з обов’язковим повторним вивченням дисципліни	не зараховано з обов’язковим повторним вивченням дисципліни

13. Методичне забезпечення

1. Комплект слайдів презентацій з матеріалами лекцій.
2. Методичні вказівки до лабораторних робіт з дисципліни “Інтелектуальний аналіз даних” для студентів напрямів підготовки 6.050103 “Програмна інженерія” та 6.050101 “Комп’ютерні науки” / Т.В. Федорончак. – Запоріжжя: ЗНТУ, 2016. – 62 с.

14. Рекомендована література

Базова

1. Барсегян А. А. Анализ данных и процессов: учеб. пособие / А.А. Барсегян, М. С. Куприянов, И. И. Холод, М. Д. Тесс, С. И. Елизаров. – 3-е изд., перераб. и доп. – СПб.: БХВ-Петербург, 2009. – 512 с.
2. Паклин Н.Б. Бизнес-аналитика: от данных к знаниям / Н.Б. Паклин, В.И. Орешков. – СПб.: Питер, 2009. – 624 с.
3. Сегаран Т. Программируем коллективный разум / Т. Сегаран. – СПб.: Символ-Плюс, 2008. – 368 с.
4. Чубукова И.А. Data Mining: учебное пособие / И.А. Чубукова. – М.: Интернет-университет информационных технологий: БИНОМ: Лаборатория знаний, 2006. – 382 с.
5. Дубровин В.И. Интеллектуальные средства диагностики и прогнозирования надежности авиадвигателей: Монография / В.И. Дубровин, С.А. Субботин, А.В. Богуслаев, В.К. Яценко. – Запорожье: ОАО "Мотор-Сич", 2003. – 279 с.

Допоміжна

6. Han J. Data Mining: Concepts and Techniques (Second Edition) / J. Han, M. Kamber – Morgan Kaufmann Publishers, 2006. – 800 p.
7. Witten, I. H. Data mining : practical machine learning tools and techniques. / Ian H. Witten, Frank Eibe, Mark A. Hall. – 3rd ed. – Morgan Kaufmann Publishers, 2011. – 630 p.
8. Макленнен Д. Microsoft SQL Server 2008: Data Mining – интеллектуальный анализ данных / Д. Макленнен, Ч. Танг, Б. Криват. – СПб.: БХВ-Петербург, 2009. – 720 с.
9. Дюк В. Data Mining : учебный курс / В. Дюк, А. Самойленко. – СПб.: Питер, 2001. – 368 с.

15. Інформаційні ресурси

1. Weka 3: Data Mining Software in Java [Електронний ресурс] – Режим доступу: http://www.cs.waikato.ac.nz/ml/weka/
2. Weka 3 Wiki documentation [Електронний ресурс] – Режим доступу: http://weka.wikispaces.com/