Форма № 3

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ**
(найменування центрального органу виконавчої влади у сфері освіти і науки)

# ІНТЕЛЕКТУАЛЬНИЙ АНАЛІЗ ДАНИХ
(назва навчальної дисципліни)

# ПРОГРАМА
**навчальної дисципліни**
**підготовки** бакалаврів
(назва рівня вищої освіти (освітньо-кваліфікаційний рівень))
**спеціальності** 121 «Інженерія програмного забезпечення»
(код і назва спеціальності)

**2017 рік**

РОЗРОБЛЕНО ТА ВНЕСЕНО: Запорізьким національним технічним університетом
(повне найменування вищого навчального закладу)

РОЗРОБНИКИ ПРОГРАМИ: Федорончак Тетяна Василівна, к.т.н., доцент кафедри ПЗ

Обговорено та рекомендовано до затвердження Вченою радою інституту, факультету, науково- методичною радою

"_____" _____ 2017 року, протокол №__

# ВСТУП

Програма вивчення нормативної навчальної дисципліни «Інтелектуальний аналіз даних» складена відповідно до освітньо-професійної програми підготовки бакалаврів спеціальності 121 «Інженерія програмного забезпечення».

**Предметом** вивчення навчальної дисципліни є технології побудови сховищ даних та методи сучасної аналітичної обробки великих об'ємів даних з метою знаходження в них нових знань,.

**Міждисциплінарні зв'язки**: Курс базується на знаннях, одержаних при вивченні дисциплін «Емпіричні методи програмної інженерії», «Математичні методи оптимізації та дослідження операцій» та «Теорія ймовірності та математична статистика».
Отриманні знання будуть використовуватися та доповнюватися при подальшому вивченні дисципліни «Теорія прийняття рішень».

Програма навчальної дисципліни складається з таких блоків змістових модулів:
1.Основи інтелектуального аналізу даних: процес виявлення знань
2.Алгоритми Data Mining: класифікація і регресія
3.Інтелектуальний аналіз часових рядів
4.Алгоритми Data Mining: кластеризація
5.Алгоритми Data Mining: пошук асоціативних правил

**1. Мета та завдання навчальної дисципліни**
1.1. Метою викладання навчальної дисципліни «Інтелектуальний аналіз даних» є вивчення методів сучасної обробки даних – інтелектуального аналізу даних (Data Mining), 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.

1.2.Основними завданнями вивчення дисципліни «Інтелектуальний аналіз даних» є:
− 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.

1.3. Згідно з вимогами освітньо-професійної програми студенти повинні:
***знати :***
− 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;
***вміти*** :
− 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.
На вивчення навчальної дисципліни відводиться 60 години/ 2 кредити ECTS.

**2. Інформаційний обсяг 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.

**3. Рекомендована література**

1. Барсегян А.А. Технологии анализа данных: Data Mining, Visual Mining, Text Mining, OLAP / А.А. Барсегян, М.С. Куприянов, В.В. Степаненко, И.И. Холод. – СПб.: БХВ-Петербург, 2007. – 384 с.

2.Олійник А. О. Інтелектуальний аналіз даних : Навчальний посібник / А. О. Олійник, О. О. Олійник, С. О. Субботін. – Запоріжжя : ЗНТУ, 2011. – 278 с.

3 Ситник В.Ф., Краснюк М.Т. Інтелектуальний аналіз даних (дейтамайнінг): навч. посібник. – К.: КНЕУ, 2007. – 376 с.

4 Флах, П. Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных / П. Флах. – М.: ДМК Пресс, 2015. – 775 с.

5.Рашка, С. Python и машинное обучение / С. Рашка. – М.: ДМК-Пресс, 2017. – 418 с.

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.—3rd ed. / Ian H. Witten, Frank Eibe, Mark A. Hall. – Morgan Kaufmann Publishers, 2011. – 629 p.

8. 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.

9. Паклин Н.Б. Бизнес-аналитика: от данных к знаниям / Н.Б. Паклин, В.И. Орешков. – СПб.: Питер, 2009. – 624 с.
10. Сегаран Т. Программируем коллективный разум / Т. Сегаран. – СПб.: Символ-Плюс, 2008. – 368 с.
11. Чубукова И.А. Data Mining: учебное пособие / И.А. Чубукова. – М.: Интернет-университет информационных технологий: БИНОМ: Лаборатория знаний, 2006. – 382 с.
12. Макленнен Д. Microsoft SQL Server 2008: Data Mining – интеллектуальный анализ данных / Д. Макленнен, Ч. Танг, Б. Криват. – СПб.: БХВ-Петербург, 2009. – 720 с.

**4. Форма підсумкового контролю успішності навчання** залік.

**5. Засоби діагностики успішності навчання** Діагностика знань студентів здійснюється з допомогою: завдань для лабораторних робіт; усних опитувань; письмових тестових та практичних робіт.