

СИЛАБУС НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

1. Загальна інформація про навчальну дисципліну

Повна назва навчальної дисципліни	Технології інтелектуального аналізу даних
Повна офіційна назва закладу вищої освіти	Сумський державний університет
Повна назва структурного підрозділу	Навчально-науковий інститут бізнес-технологій "УАБС". Кафедра економічної кібернетики
Розробник(и)	Гриценко Костянтин Григорович
Рівень вищої освіти	Перший рівень вищої освіти, НРК – 6 рівень, QF-LLL – 6 рівень, FQ-EHEA – перший цикл
Семестр вивчення навчальної дисципліни	16 тижнів протягом 4-го семестру
Обсяг навчальної дисципліни	Обсяг дисципліни становить 5 кред. ЄКТС, 150 год., з яких 48 год. становить контактна робота з викладачем (16 год. лекцій, 32 год. лабораторних занять)
Мова викладання	Українська, англійська

2. Місце навчальної дисципліни в освітній програмі

Статус дисципліни	Вибіркова навчальна дисципліна для освітніх програм "Економіка природокористування", "Економіка і бізнес", "Бізнес аналітика", "Економіка та бізнес-інновації", "Економічна кібернетика"
Передумови для вивчення дисципліни	Передумови для вивчення відсутні
Додаткові умови	Додаткові умови відсутні
Обмеження	Обмеження відсутні

3. Мета навчальної дисципліни

Метою навчальної дисципліни є формування у студентів професійних компетенцій щодо підготовки даних для аналізу, володіння методами інтелектуального аналізу даних, обґрунтованого вибору технології інтелектуального аналізу даних, визначення чинників, що впливають на результати аналізу даних, прогнозування найімовірнішого перебігу економічних процесів на перспективу, використання відповідних критеріїв для аналізу рівня достовірності отриманих оцінок

4. Зміст навчальної дисципліни

Тема 1 Основи інтелектуального аналізу даних Поняття, суть, ціль та технологія інтелектуального аналізу даних. Розвиток і призначення дейтамайнінгу. Інтелектуальний аналіз даних як частина ринку інформаційних технологій. Інтелектуальний аналіз даних як мультидисциплінарна область. Задачі, методи та алгоритми інтелектуального аналізу даних. Характеристика процесів і активностей дейтамайнінгу. Проблеми використання технології інтелектуального аналізу даних.
--

Тема 2 Вступ до SAS Enterprise Miner

Основні етапи та алгоритми інтелектуального аналізу даних. Огляд бізнес-аналітичних систем. Огляд алгоритмів і інформаційних систем дейтамайнінгу. Огляд бізнес-аналітичних рішень SAS Institute Inc. Огляд інтерфейсу SAS Enterprise Miner. Аналітичний процес в SAS Enterprise Miner. Огляд та використання технології SEMMA.

Тема 3 Первинний аналіз даних у SAS Enterprise Miner

Архітектура і класифікація систем дейтамайнінгу. Стратегія вибору та оцінювання систем дейтамайнінгу. Характеристика сучасного стану ринку систем дейтамайнінгу та напрями розвитку. Клієнт-серверна архітектура SAS Enterprise Miner. Організація елементів аналізу в SAS Enterprise Miner. Створення проекту та діаграми SAS Enterprise Miner. Завдання джерела даних. Аналіз вхідних даних. Налаштування даних, ролей та вимірів змінних. Основи data quality. Змінювання настроювань вибірок по замовчуванню. Аналіз змінних джерела даних. Створення діаграми для фільтрації непотрібних елементів.

Тема 4 Кластерний аналіз у SAS Enterprise Miner

Основи виявлення структури даних. Постановка задачі кластеризації. Міри відстані або схожості, що використовуються в задачах кластеризації. Матриця відстаней (схожості). Дерево класифікації алгоритмів кластеризації. Ієрархічні алгоритми кластеризації. Неієрархічні алгоритми кластеризації. Представлення результатів кластеризації. Метод «найближчого сусіда» та «k-найближчого сусіда». Безнаглядна класифікація. Інструменти виявлення структури даних в SAS Enterprise Miner. Особливості реалізації алгоритму кластеризації методом k-середніх в SAS Enterprise Miner. Основні етапи проведення кластерного аналізу в SAS Enterprise Miner: попередній аналіз вхідних даних, підготовка даних до проведення кластерного аналізу, визначення оптимального числа кластерів і проведення процедури кластерного аналізу, профілювання сегментів і інтерпретація результатів.

Тема 5 Асоціативний аналіз у SAS Enterprise Miner

Різновиди задачі пошуку асоціативних правил. Представлення результатів асоціативного аналізу. Алгоритм Apriori та його різновиди. Введення в аналіз ринкового кошика. Проведення аналізу ринкового кошика засобами SAS Enterprise Miner. Інтерпретація одержаних результатів і розробка практичних рекомендацій. Постановка задачі пошуку асоціативних правил. Секвенціальний аналіз.

Тема 6 Прогнозне моделювання із використанням дерев рішень у SAS Enterprise Miner

Типи прогнозів: рішення, рейтинги та оцінки. Дерева рішень як метод дейтамайнінгу. Дерева рішень для дослідження альтернатив. Дерева рішень для класифікації. Інструменти прогнозного моделювання в SAS Enterprise Miner. Набори даних для прогнозної моделі: навчальні, перевірочні та тестові. Проблема розмірності. Інтерактивне вирощування дерев рішень засобами SAS Enterprise Miner. Оцінка прогнозної моделі з деревом рішень в SAS Enterprise Miner. Задачі скорочення та автономного вирощування дерев рішень. Засоби інтерпретації та оцінки дерева рішень в SAS Enterprise Miner.

Тема 7 Прогнозне моделювання із використанням регресійного аналізу в SAS Enterprise Miner

Прогнозна формула лінійної та логістичної регресії. Логіт-функція та логістична функція. Оцінки параметрів регресії з використанням методу максимальної правдоподібності. Обробка пропущених значень в SAS Enterprise Miner. Побудова моделей регресійного типу в SAS Enterprise Miner. Вибір вхідних змінних регресії. Оптимізація складності структури регресійного рівняння. Інтерпретація регресійних моделей. Трансформація вхідних змінних в SAS Enterprise Miner. Робота з вхідними змінними категоріального типу. Поліноміальна регресійна модель.

Тема 8 Прогнозне моделювання із використанням нейронних мереж у SAS Enterprise Miner

Математична модель штучного нейрона. Архітектура нейромереж. Загальний опис навчання та використання нейромереж. Приклади вдалого використання нейромереж у бізнесі. Переваги й обмеження нейромереж. Прогнозна формула нейронної мережі. Прогнозні оцінки. Вибір вхідних змінних нейронної мережі. Інструменти навчання нейронної мережі в SAS Enterprise Miner. Критерії зупинки навчання нейронної мережі.

5. Очікувані результати навчання навчальної дисципліни

Після успішного вивчення навчальної дисципліни здобувач вищої освіти зможе:

PH1	знати концептуальні положення інтелектуального аналізу даних
PH2	застосовувати в професійній діяльності технології інтелектуального аналізу даних
PH3	знати та володіти сучасним інструментарієм інтелектуального аналізу даних

7. Види навчальних занять та навчальної діяльності

7.1 Види навчальних занять

Тема 1. Основи інтелектуального аналізу даних
Лк1 "Основи інтелектуального аналізу даних" (денна) Поняття, суть, ціль та технологія інтелектуального аналізу даних. Розвиток і призначення дейтамайнінгу. Інтелектуальний аналіз даних як частина ринку інформаційних технологій. Інтелектуальний аналіз даних як мультидисциплінарна область. Задачі, методи та алгоритми інтелектуального аналізу даних. Характеристика процесів і активностей дейтамайнінгу. Проблеми використання технології інтелектуального аналізу даних.
Тема 2. Вступ до SAS Enterprise Miner
Лк2 "Вступ до SAS Enterprise Miner" (денна) Основні етапи та алгоритми інтелектуального аналізу даних. Огляд бізнес-аналітичних систем. Огляд алгоритмів і інформаційних систем дейтамайнінгу. Огляд бізнес-аналітичних рішень SAS Institute Inc. Огляд інтерфейсу SAS Enterprise Miner. Аналітичний процес в SAS Enterprise Miner. Огляд та використання технології SEMMA.
Тема 3. Первинний аналіз даних у SAS Enterprise Miner
Лк3 "Первинний аналіз даних у SAS Enterprise Miner" (денна) Архітектура і класифікація систем дейтамайнінгу. Стратегія вибору та оцінювання систем дейтамайнінгу. Характеристика сучасного стану ринку систем дейтамайнінгу та напрями розвитку. Клієнт-серверна архітектура SAS Enterprise Miner. Організація елементів аналізу в SAS Enterprise Miner. Створення проекту та діаграми SAS Enterprise Miner. Завдання джерела даних. Аналіз вхідних даних. Налаштування даних, ролей та вимірів змінних. Основи data quality. Змінювання настроювань вибірок по замовчуванню. Аналіз змінних джерела даних. Створення діаграми для фільтрації непотрібних елементів.
Тема 4. Кластерний аналіз у SAS Enterprise Miner

Лк4 "Кластерний аналіз у SAS Enterprise Miner" (денна)

Основи виявлення структури даних. Постановка задачі кластеризації. Міри відстані або схожості, що використовуються в задачах кластеризації. Матриця відстаней (схожості). Дерево класифікації алгоритмів кластеризації. Ієрархічні алгоритми кластеризації. Неієрархічні алгоритми кластеризації. Представлення результатів кластеризації. Метод «найближчого сусіда» та «k-найближчого сусіда». Безнаглядна класифікація. Інструменти виявлення структури даних в SAS Enterprise Miner. Особливості реалізації алгоритму кластеризації методом k-середніх в SAS Enterprise Miner. Основні етапи проведення кластерного аналізу в SAS Enterprise Miner: попередній аналіз вхідних даних, підготовка даних до проведення кластерного аналізу, визначення оптимального числа кластерів і проведення процедури кластерного аналізу, профілювання сегментів і інтерпретація результатів.

Лб1 "Кластерний аналіз в SAS Enterprise Miner" (денна)

Оволодіння базовими навиками роботи з системою SAS Enterprise Miner

Лб2 "Кластерний аналіз в SAS Enterprise Miner" (денна)

Оволодіння базовими навиками проведення первинного аналізу якості даних з використанням SAS Enterprise Miner

Лб3 "Кластерний аналіз в SAS Enterprise Miner" (денна)

Оволодіння інструментами виявлення структури даних із використанням SAS Enterprise Miner

Лб4 "Кластерний аналіз в SAS Enterprise Miner" (денна)

Освоєння технології кластерного аналізу з використанням SAS Enterprise Miner

Тема 5. Асоціативний аналіз у SAS Enterprise Miner

Лк5 "Асоціативний аналіз у SAS Enterprise Miner" (денна)

Різновиди задачі пошуку асоціативних правил. Представлення результатів асоціативного аналізу. Алгоритм Аргіогі та його різновиди. Введення в аналіз ринкового кошика. Проведення аналізу ринкового кошика засобами SAS Enterprise Miner. Інтерпретація одержаних результатів і розробка практичних рекомендацій. Постановка задачі пошуку асоціативних правил. Секвенціальний аналіз.

Лб5 "Асоціативний аналіз у SAS Enterprise Miner" (денна)

Особливості проведення аналізу ринкового кошика засобами SAS Enterprise Miner. Постановка задачі пошуку асоціативних правил. Різновиди задачі пошуку асоціативних правил. Освоєння технології асоціативного аналізу з використанням SAS Enterprise Miner

Лб6 "Асоціативний аналіз у SAS Enterprise Miner" (денна)

Освоєння технології секвенціального аналізу з використанням SAS Enterprise Miner

Тема 6. Прогнозне моделювання із використанням дерев рішень у SAS Enterprise Miner

Лк6 "Прогнозне моделювання із використанням дерев рішень у SAS Enterprise Miner" (денна)

Типи прогнозів: рішення, рейтинги та оцінки. Дерева рішень як метод дейтамайнінгу. Дерева рішень для дослідження альтернатив. Дерева рішень для класифікації. Інструменти прогнозного моделювання в SAS Enterprise Miner. Набори даних для прогновної моделі: навчальні, перевірочні та тестові. Проблема розмірності. Інтерактивне вирощування дерев рішень засобами SAS Enterprise Miner. Оцінка прогновної моделі з деревом рішень в SAS Enterprise Miner. Задачі скорочення та автономного вирощування дерев рішень. Засоби інтерпретації та оцінки дерева рішень в SAS Enterprise Miner.

Лб7 "Прогнозне моделювання із використанням дерев рішень у SAS Enterprise Miner" (денна)

Набори даних для прогновної моделі: навчальні, перевірочні та тестові. Проблема розмірності. Інтерактивне вирощування дерев рішень засобами SAS Enterprise Miner.

Лб8 "Прогнозне моделювання із використанням дерев рішень у SAS Enterprise Miner" (денна)

Засоби інтерпретації та оцінки дерева рішень в SAS Enterprise Miner. Оцінка побудованого дерева рішень у SAS Enterprise Miner.

Лб9 "Прогнозне моделювання із використанням дерев рішень у SAS Enterprise Miner" (денна)

Задачі скорочення та автономного вирощування дерев рішень. Автоматичне вирощування дерева рішень у SAS Enterprise Miner.

Тема 7. Прогнозне моделювання із використанням регресійного аналізу в SAS Enterprise Miner

Лк7 "Прогнозне моделювання із використанням регресійного аналізу в SAS Enterprise Miner" (денна)

Прогнозна формула лінійної та логістичної регресії. Логіт-функція та логістична функція. Оцінки параметрів регресії з використанням методу максимальної правдоподібності. Обробка пропущених значень в SAS Enterprise Miner. Побудова моделей регресійного типу в SAS Enterprise Miner. Вибір вхідних змінних регресії. Оптимізація складності структури регресійного рівняння. Інтерпретація регресійних моделей. Трансформація вхідних змінних в SAS Enterprise Miner. Робота з вхідними змінними категоріального типу. Поліноміальна регресійна модель.

Лб10 "Прогнозне моделювання із використанням регресійного аналізу в SAS Enterprise Miner" (денна)

Освоєння технології прогнозного моделювання із використанням регресійного аналізу в SAS Enterprise Miner. Обробка пропущених значень в SAS Enterprise Miner. Вибір вхідних змінних регресії. Трансформація вхідних змінних в SAS Enterprise Miner. Робота з вхідними змінними категоріального типу.

Лб11 "Прогнозне моделювання із використанням регресійного аналізу в SAS Enterprise Miner" (денна)

Побудова моделей регресійного типу в SAS Enterprise Miner. Оптимізація складності структури регресійного рівняння. Інтерпретація регресійних моделей.

Лб12 "Прогнозне моделювання із використанням регресійного аналізу в SAS Enterprise Miner" (денна)

Прогнозне моделювання із використанням поліноміальної регресії в SAS Enterprise Miner

Лб13 "Скорингові моделі в SAS Enterprise Miner" (денна) Освоєння методики побудови скорингових моделей в SAS Enterprise Miner. Алгоритм логістичної регресії. Створення скорингового джерела даних в SAS Enterprise Miner.
Лб14 "Скорингові моделі в SAS Enterprise Miner" (денна) Робота з компонентами ScoreCard, Interactive Grouping, Reject Interface. Експорт скорингової таблиці. Рішення SAS Credit Scoring.
Тема 8. Прогнозне моделювання із використанням нейронних мереж у SAS Enterprise Miner
Лк8 "Прогнозне моделювання із використанням нейронних мереж у SAS Enterprise Miner" (денна) Математична модель штучного нейрона. Архітектура нейромереж. Загальний опис навчання та використання нейромереж. Приклади вдалого використання нейромереж у бізнесі. Переваги й обмеження нейромереж. Прогнозна формула нейронної мережі. Прогнозні оцінки. Вибір вхідних змінних нейронної мережі. Інструменти навчання нейронної мережі в SAS Enterprise Miner. Критерії зупинки навчання нейронної мережі.
Лб15 "Прогнозне моделювання із використанням нейронних мереж у SAS Enterprise Miner" (денна) Освоєння технології прогнозного моделювання із використанням нейронних мереж у SAS Enterprise Miner. Вибір вхідних змінних нейронної мережі. Оволодіння інструментами навчання нейронної мережі в SAS Enterprise Miner. Критерії зупинки навчання нейронної мережі.
Лб16 "Прогнозне моделювання із використанням нейронних мереж у SAS Enterprise Miner" (денна) Освоєння технології прогнозного моделювання із використанням нейронних мереж у SAS Enterprise Miner. Робота з вхідними змінними категоріального типу.

7.2 Види навчальної діяльності

НД1	Виконання та презентація результатів лабораторної роботи
НД2	Самостійна підготовка до захисту лабораторної роботи
НД3	Захист звіту лабораторної роботи
НД4	Підготовка до поточного та підсумкового контролю

8. Методи викладання, навчання

Дисципліна передбачає навчання через:

МН1	Інтерактивні, тематичні, оглядові лекції
МН2	Лабораторні заняття

Лекції надають студентам матеріали з методів, алгоритмів і технологій інтелектуального аналізу даних у SAS Enterprise Miner (PH1). Лекції доповнюються лабораторними заняттями, що надають студентам можливість застосовувати теоретичні знання на практичних прикладах (PH1, PH2, PH3). Зміст лабораторних робіт спрямований на практико-орієнтоване навчання, що передбачає застосування студентами відповідних технологій інтелектуального аналізу даних для вирішення економічних завдань, аналізу інформації та підготовки аналітичних звітів (PH 2, PH 3).

Самостійному навчанню сприятиме підготовка до лекцій і лабораторних занять. Під час виконання лабораторних занять студенти розвиватимуть навички самостійного навчання, здатність до використання сучасних програмних платформ реалізації алгоритмів, економіко-математичних та

імітаційних моделей. Необхідним елементом успішного засвоєння матеріалу навчальної дисципліни є самостійна робота студентів з вітчизняною та закордонною спеціальною літературою, спеціальними пакетами аналізу статистичних даних та моделювання економічних процесів та систем, періодичними виданнями.

9. Методи та критерії оцінювання

9.1. Критерії оцінювання

Шкала оцінювання ECTS	Визначення	Чотирибальна національна шкала оцінювання	Рейтингова бальна шкала оцінювання
A	Відмінне виконання лише з незначною кількістю помилок	5 (відмінно)	$90 \leq RD \leq 100$
B	Вище середнього рівня з кількома помилками	4 (добре)	$82 \leq RD < 89$
C	Загалом правильна робота з певною кількістю помилок	4 (добре)	$74 \leq RD < 81$
D	Непогано, але зі значною кількістю недоліків	3 (задовільно)	$64 \leq RD < 73$
E	Виконання задовольняє мінімальні критерії	3 (задовільно)	$60 \leq RD < 63$
FX	Можливе повторне складання	2 (незадовільно)	$35 \leq RD < 59$
F	Необхідний повторний курс з навчальної дисципліни	2 (незадовільно)	$0 \leq RD < 34$

9.2 Методи поточного формативного оцінювання

МФО1	Опитування та усні коментарі викладача за його результатами
МФО2	Обговорення та самокорекція виконаної роботи студентами
МФО3	Настанови викладача в процесі виконання лабораторних завдань

9.3 Методи підсумкового сумативного оцінювання

МСО1	Звіт за результатами виконання лабораторних робіт
МСО2	Поточні контрольні роботи (проміжний модульний контроль)

Контрольні заходи:

4 семестр		100 балів
МСО1. Звіт за результатами виконання лабораторних робіт		90
	6x15	90
МСО2. Поточні контрольні роботи (проміжний модульний контроль)		10
	2x5	10

Контрольні заходи в особливому випадку:

4 семестр		100 балів
МСО1. Звіт за результатами виконання лабораторних робіт		90

	виконання практичного кейсу	90
МСО2. Поточні контрольні роботи (проміжний модульний контроль)		10
	із використанням тестових технологій (2x5)	10

Оцінювання знань студента під час лабораторних занять має на меті перевірку рівня підготовленості студента до виконання конкретної роботи, для чого студент в аудиторії проходить опитування (МФО1), обговорення та самокорекцію виконаної роботи (МФО2), контроль активності під час виконання роботи (МФО3).

10. Ресурсне забезпечення навчальної дисципліни

10.1 Засоби навчання

ЗН1	Програмний пакет SAS Enterprise Miner
ЗН2	Інформаційно-комунікаційні системи
ЗН3	Бібліотечні фонди

10.2 Інформаційне та навчально-методичне забезпечення

Основна література	
1	Jim Georges, Jeff Thompson, Chip Wells. Applied analytics using SAS Enterprise Miner Course Notes. – Cary: SAS Institute Inc., 2017. – 694 p.
2	Петер Флах. Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных. – М.: ДМК ПРЕСС, 2015. – 399 с.
Допоміжна література	
1	Черняк О.І. Інтелектуальний аналіз даних: підручник / О.І. Черняк, П.В. Захарченко. – К.: Знання, 2014. – 599 с.
2	Інтелектуальний аналіз даних: практикум / М.Т. Фісун, І.О. Кравець, П.П. Казмірчук, С.Г. Ніколенко. – Львів: Новий Світ-2000, 2019. – 162 с.
3	Інтелектуальний аналіз даних: комп'ютерний практикум / О.О. Сергеев-Горчинський, Г.В. Іщенко. – Київ: КПІ ім. Ігоря Сікорського, 2018. – 73 с.
Інформаційні ресурси в Інтернеті	
1	Інформатика та математичні методи в моделюванні. – Режим доступу: http://immm.opu.ua/
2	Математичне моделювання в економіці. – Режим доступу: http://www.icyb.kiev.ua/s/396/ua/
3	Системні дослідження та інформаційні технології. – Режим доступу: http://journal.iasa.kpi.ua/