

СИЛАБУС ОСВІТНЬОЇ КОМПОНЕНТИ



Інтелектуальний аналіз даних (Data Mining)

спеціальність	Не обмежено	обов'язковість дисципліни	вибіркова
освітня програма	Не обмежено	факультет	Інститут "Кіберпорт"
освітній рівень	перший (бакалаврський)	кафедра	Кафедра кібернетики та інформаційних технологій

ВИКЛАДАЧ

Міхнова Олена Дмитрівна



Вища освіта – Харківський національний університет радіоелектроніки, спеціальність «Інформаційні управляючі системи та технології»; Харківський гуманітарний університет «Народна українська академія», спеціальність «Переклад»

Науковий ступень – к.т.н.

Вчене звання – доцент

Досвід роботи – більше 5 років

Показники професійної активності з тематики курсу:

- співавторка більше 10 методичних розробок;
- співавторка 2 навчальних посібників;
- учасниця наукових і методичних конференцій.

телефон	+380955053180	електронна пошта	mikhnova@btu.kharkov.ua	дистанційна підтримка	Google Meet Moodle, ZOOM
---------	---------------	------------------	--	-----------------------	-----------------------------

ЗАГАЛЬНА ІНФОРМАЦІЯ ПРО ОСВІТНЮ КОМПОНЕНТУ (ДИСЦИПЛІНУ)

Мета	формування системи знань про типи завдань, що виникають в інтелектуальному аналізі даних (Data Mining), здатності до аналізу великих масивів інформації з метою виявлення нових знань і вмінь, необхідних для прийняття рішень; вивчення основних методів і моделей сучасної обробки даних; формування практичних навичок роботи із пакетами прикладних програм для розв'язання задач аналізу та інтерпретації даних, розгляд практичних прикладів застосування Data Mining; підготовка студентів до самостійної роботи з вирішення задач засобами Data Mining і розробки інтелектуальних систем
Формат	лекції, практичні заняття, самостійна робота, командна робота
Деталізація результатів навчання і форм їх контролю	<ul style="list-style-type: none">• Вміти обґрунтовувати вибір конкретного типу моделі та методу інтелектуального аналізу даних при вирішенні поставленої практичної задачі / індивідуальні практичні завдання• 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 / індивідуальні практичні завдання• Вміти використовувати сучасні програмні засоби для проектування та дослідження систем інтелектуального аналізу даних / індивідуальні практичні завдання• Мати навички із використання методів прогнозування, аналізу багатомірних угруповань, статистичної обробки часових рядів і прогнозування, класифікації об'єктів у випадку невідомих розподілів даних / індивідуальні практичні завдання
Обсяг і форми контролю	3 кредити ECTS (90 годин): 12 годин лекції, 18 годин лабораторно-практичні; модульний контроль (2 модулі); підсумковий контроль – диференційований залік.
Вимоги викладача	вчасне виконання завдань, активність, командна робота
Умови зарахування на освітню компоненту	вільне зарахування

СТРУКТУРА ОСВІТНЬОЇ КОМПОНЕНТИ (ДИСЦИПЛІНИ)

Модуль 1. Статистичні методи інтелектуального аналізу

Лекція 1.	Основні поняття та визначення інтелектуального аналізу даних та математичної статистики. Суть, мета та сфера застосування технології Data Mining.	Лабораторно-практичне заняття 1 (ЛПЗ 1)	Визначення числових характеристик вибірки та побудова графіків.	Самостійна робота	1. Системи підтримки прийняття рішень. Задачі систем підтримки прийняття рішень. 2. Бази, банки та сховища даних. Концепції та організація сховищ даних. 3. OLAP-технології. 4. Попередня обробка даних, опрацювання стурктурованих та не стурктурованих даних.
Лекція 2.	Перевірка статистичних гіпотез	ЛПЗ 2	Перевірка статистичних гіпотез.		
Лекція 3.	Кореляційний і регресійний аналіз даних. Множинний регресійний аналіз. Лінійна множинна регресійна модель. Перевірка адекватності моделі.	ЛПЗ 3	Кореляційний аналіз даних.		
		ЛПЗ 4	Побудова регресійних моделей. Перевірка адекватності моделі.		

Модуль 2. Методи машинного навчання

Лекція 4.	Задачі прогнозування. Аналіз часових рядів. Тренд, сезонність і цикл. Методи візуалізації Data Mining.	ЛПЗ 5	Прогнозування і часові ряди. Види помилок та прогнозів.	Самостійна робота	5. Нейромережеві технології інтелектуального аналізу даних. Архітектура нейронних мереж. 6. Нечіткі методи інтелектуального аналізу даних. Концепція нечітких обчислень. 7. CRISP-DM - міжгалузевий стандарт процесу аналізу даних і підтримки прийняття рішень. 8. Генетичні алгоритми. 9. Онтології та технології text-mining.
Лекція 5.	Методи кластеризації та класифікації. Види кластерів. Міри близькості, засновані на відстанях. Базові алгоритми кластеризації.	ЛПЗ 6	Методи кластерного аналізу. Міри близькості, засновані на відстанях.		
		ЛПЗ 7	Методи класифікації. Постановка задачі класифікації та представлення результатів.		
Лекція 6.	Основні поняття теорії асоціативних правил та дерева рішень.	ЛПЗ 8	Дерева рішень. Методи опорних векторів, «найближчого сусіда», Баеса.		
		ЛПЗ 9	Пошук асоціативних правил на основі наявних даних.		

ОСНОВНА ЛІТЕРАТУРА ТА МЕТОДИЧНІ МАТЕРІАЛИ

Література

1. Акіменко В.В. Прикладні задачі інтелектуального аналізу даних (DATA MINING). К.: КНУ ім. Тараса Шевченка, 2018. 152 с.
2. Черняк О.І., Захарченко П.В. Інтелектуальний аналіз даних: підручник. К., 2014. 599 с.
3. Снитюк В. Є. Прогнозування. Моделі. Методи. Алгоритми : навчальний посібник. К.: Маклаут, 2008. 364 с.
4. Чорней Р. К. Практикум з теорії ймовірностей та математичної статистики. К., 2006. 328 с.
4. Майборода Р.Є. "Комп'ютерна статистика". ВПЦ "Київський університет", 2019. - 589 с.
5. Майборода Р.Є., Сугакова О.В. Аналіз даних за допомогою пакета R. Навчальний посібник. Київ, Київський національний університет ім. Тараса Шевченка, 2015.- 65 с. В. Гнатюк. Вступ до R на прикладах. Харків: Харківський національний економічний університет, 2010. – 101 с.
6. A. Agresti. Statistical methods for social sciences. Boston: Pearson, 2018.
7. Zumel N., Mount J. Practical Data Science with R. - Manning Publications Co., 2014. – 417 p.

Методичне забезпечення

1. Василенко О. А. Математично-статистичні методи аналізу у прикладних дослідженнях: навч. посіб. / О. А. Василенко, І. А. Сенча. – Одеса: ОНАЗ ім. О. С. Попова, 2011. – 166 с.
2. О. Сергеев-Горчинський, Г. В. Іщенко ; КПІ ім. Ігоря Сікорського. – К.: КПІ ім. Ігоря Сікорського, 2018. – 73 с.
3. Мамчич Т. І., Оленко А.Я., Осипчук М.М., Шпортюк В.Г. Статистичний аналіз даних з пакетом Statistica. Дрогобич: ВФ „Відродження”, 2006. – 208с.
4. Олійник А. О., Субботін С. О., Олійник О. О.. Інтелектуальний аналіз даних : навчальний посібник, Запоріжжя : ЗНТУ, 2012, 278 с.
5. Ситник В. Ф., Краснюк М.Т. Інтелектуальний аналіз даних (дейтамайнінг): Навч. посібник. К: КНЕУ, 2007. 376 с.

СИСТЕМА ОЦІНЮВАННЯ

СИСТЕМА		БАЛИ	ДІЯЛЬНІСТЬ, ЩО ОЦІНЮЄТЬСЯ
Підсумкове оцінювання	100 бальна ECTS (стандартна)	до 50	50% від усередненої оцінки за модулі
		до 50	підсумкове тестування
Модульне оцінювання	100 бальна сумарна	до 50	відповіді на тестові питання
		до 20	усні відповіді на лабораторно-практичних заняттях
		до 30	результат засвоєння блоку самостійної роботи

НОРМИ АКАДЕМІЧНОЇ ЕТИКИ ТА ДОБРОЧЕСНОСТІ

Всі учасники освітнього процесу (в тому числі здобувачі освіти) повинні дотримуватися кодексу академічної доброчесності та вимог, які прописані у положенні «Про академічну доброчесність учасників освітнього процесу ДБТУ»: виявляти дисциплінованість, вихованість, поважати гідність один одного, проявляти доброзичливість, чесність, відповідальність.