

ВИЩИЙ НАВЧАЛЬНИЙ ЗАКЛАД УКООПСІЛКИ  
 «ПОЛТАВСЬКИЙ УНІВЕРСИТЕТ ЕКОНОМІКИ І ТОРГІВЛІ»  
 Навчально-науковий інститут денної освіти  
 Кафедра комп'ютерних наук та інформаційних технологій

## СИЛАБУС

навчальної дисципліни

### «Аналіз даних та прикладні пакети статистичної обробки»

на 2021-2022 навчальний рік

Курс та семестр вивчення	4 курс, 1 семестр
Освітня програма/спеціалізація	122 Комп'ютерні науки
Спеціальність	122 Комп'ютерні науки
Галузь знань	12 «Інформаційні технології»
Ступінь вищої освіти	бакалавр

ПІБ НПП, який веде дану дисципліну,  
 науковий ступінь і вчене звання,  
 посада

Олексійчук Ю.Ф., доцент кафедри комп'ютерних наук та інформаційних технологій, к.ф.-м.н.

Контактний телефон	0532 509204
Електронна адреса	<a href="mailto:olexijchuk@gmail.com">olexijchuk@gmail.com</a>
Розклад навчальних занять	<a href="http://schedule.puet.edu.ua/">http://schedule.puet.edu.ua/</a>
Консультації	он-лайн: електронною поштою, за розкладом <a href="http://www.matmodel.puet.edu.ua">www.matmodel.puet.edu.ua</a> вкладка Студентів
Сторінка дистанційного курсу	<a href="http://www2.el.puet.edu.ua/">http://www2.el.puet.edu.ua/</a>

#### Опис навчальної дисципліни

<b>Мета вивчення навчальної дисципліни</b>	Основною метою навчальної дисципліни «Аналіз даних та прикладні пакети статистичної обробки» є формування уявлення про методи статистичної обробки даних із використанням комп'ютерних технологій для розв'язування практичних задач; виробити навички вибору відповідних статистичних моделей та методів для наявних даних.
<b>Тривалість</b>	4 кредити ЄКТС/120 годин (лекції – 16 год. практичні роботи 32 год., самостійна робота 72 год.)
<b>Форми та методи навчання</b>	практичні заняття в аудиторії, самостійна робота поза розкладом; словесні - розповідь, пояснення; наочні - демонстрація та ілюстрація; практичні - практична робота.
<b>Система поточного та підсумкового контролю</b>	відвідування занять; виконання практичних робіт, МКР Підсумковий контроль: ПМК
<b>Базові знання</b>	Курс базується на таких дисциплінах: Теорія ймовірностей і математична статистика
<b>Мова викладання</b>	Українська

#### Перелік компетентностей, які забезпечує дана навчальна дисципліна, програмні результати навчання

Програмні результати навчання	Компетентності, якими повинен оволодіти здобувач
ПР1. Застосовувати знання основних форм і законів абстрактно-логічного мислення, основ методології наукового пізнання, форм і методів вилучення, аналізу, обробки та синтезу інформації в предметній області комп'ютерних наук.	Здатність до абстрактного мислення, аналізу та синтезу (ЗК1). Здатність застосовувати знання у практичних ситуаціях (ЗК2). Знання та розуміння предметної області та розуміння професійної діяльності (ЗК3).
ПР3. Використовувати знання закономірностей випадкових явищ, їх властивостей та операцій над ними, моделей випадкових процесів та сучасних програмних середовищ для розв'язування задач статистичної обробки даних і побудови прогнозних моделей.	Здатність спілкуватися державною мовою як усно, так і письмово (ЗК4). Здатність вчитися й оволодівати сучасними знаннями (ЗК6). Здатність до пошуку, оброблення та аналізу інформації з різних джерел (ЗК7). Здатність генерувати нові ідеї (креативність) (ЗК8). Здатність працювати в команді (ЗК9).

<p>ПР12. Застосовувати методи та алгоритми обчислювального інтелекту та інтелектуального аналізу даних в задачах класифікації, прогнозування, кластерного аналізу, пошуку асоціативних правил з використанням програмних інструментів підтримки багатовимірних аналізів даних на основі технологій DataMining, TextMining, WebMining..</p>	<p>Здатність бути критичним і самокритичним (ЗК10).  Здатність приймати обґрунтовані рішення (ЗК11)  Здатність оцінювати та забезпечувати якість виконуваних робіт (ЗК12).  Здатність діяти на основі етичних міркувань (ЗК13).  Здатність до виявлення статистичних закономірностей недетермінованих явищ, застосування методів обчислювального інтелекту, зокрема статистичної, нейромережевої та нечіткої обробки даних, методів машинного навчання та генетичного програмування тощо (СК2).  Здатність застосовувати теоретичні та практичні основи методології та технології моделювання для дослідження характеристик і поведінки складних об'єктів систем, проводити обчислювальні експерименти з обробкою й аналізом результатів (СК7).  Здатність до інтелектуального аналізу даних на основі методів обчислювального інтелекту включно з великими та погано структурованими даними, їхньої оперативної обробки та візуалізації результатів аналізу в процесі розв'язування прикладних задач (СК11).</p>
--	---

### **Тематичний план навчальної дисципліни**

1 Назва теми (лекції) та питання теми (лекції)	2 Назва теми та питання семінарського, практичного або лабораторного заняття	3 Завдання самостійної роботи в розрізі тем
<b>Модуль 1. Основи теорії програмування. Контекстовільні мови. Синтаксичний аналіз.</b>		
<p><b>Тема 1. Попередній аналіз</b>  <i>Лекція 1. Підготовка вхідної інформації. Попередня обробка. Аномальні спостереження. Розвідувальний аналіз. Гіпотези та їх перевірка</i></p>	<p><u>Практичне заняття 1.</u> Знайомство з Anaconda та Python  <u>Практичне заняття 2.</u> Видалення аномальних спостережень  <u>Практичне заняття 3.</u> Робота з даними в Python</p>	<p>опрацювати лекційний матеріал до теми 1, готуватись до практичних занять, опрацювати ДК</p>
<p><b>Тема 2. Кореляційний аналіз</b>  <i>Лекція 2. Кореляційний аналіз. Кореляція якісних змінних</i></p>	<p><u>Практичне заняття 4.</u> Перевірка стохастичності вибірки  <u>Практичне заняття 5.</u> Основні статистичні характеристики</p>	<p>опрацювати лекційний матеріал до теми 2, готуватись до практичних занять, опрацювати ДК</p>
<p><b>Тема 3. Регресійний аналіз</b>  <i>Лекція 3. Регресійний аналіз</i></p>	<p><u>Практичне заняття 6.</u> Коефіцієнт кореляції  <u>Практичне заняття 7.</u> Кореляція якісних змінних  <u>Практичне заняття 8.</u> Частинна кореляція</p>	<p>опрацювати лекційний матеріал до теми 3, готуватись до практичних занять, опрацювати ДК</p>
<p><b>Тема 4. Дисперсійний аналіз</b>  <i>Лекція 4. Дисперсійний аналіз</i></p>	<p><u>Практичне заняття 9.</u>  Парна лінійна регресія</p>	
<p><b>Тема 5. Дискримінантний аналіз</b></p>	<p><u>Практичне заняття 10.</u> Обчислення параметрів парної регресії засобами MS Excel</p>	<p>опрацювати лекційний матеріал до теми 4, готуватись до практичних занять, опрацювати</p>

1	2	3
<p><u>Лекція 5. Дискримінантний аналіз</u></p> <p><b>Тема 6. Кластерний аналіз</b></p> <p><u>Лекція 6. Кластерний аналіз</u></p>	<p>Практичне заняття 11. Парна нелінійна регресія</p> <p>Практичне заняття 12. Множинна регресія</p> <p>Практичне заняття 13. Модульна контрольна робота №1</p>	<p>ДК</p> <p>опрацювати лекційний матеріал до теми 5, готуватись до практичних занять, опрацювати ДК</p> <p>опрацювати лекційний матеріал до теми 6, готуватись до практичних занять, опрацювати ДК, готуватись до модульної контрольної роботи №1</p>
<b>Модуль 2. Аналіз даних з використанням статистичних пакетів</b>		
<p><b>Тема 7. Аналіз двовимірних даних</b></p> <p><u>Лекція 7. Аналіз двовимірних даних</u></p> <p><b>Тема 8. Класифікація даних</b></p> <p><u>Лекція 8. Класифікація даних</u></p>	<p>Практичне заняття 14 Метод k-найближчих сусідів.</p> <p>Практичне заняття 15 Модульна контрольна робота № 2.</p>	<p>опрацювати лекційний матеріал до теми 7, готуватись до практичних занять, опрацювати ДК</p> <p>опрацювати лекційний матеріал до теми 8, готуватись до практичних занять, опрацювати ДК, готуватись до модульної контрольної роботи №2</p>

### **Інформаційні джерела**

1. Фоусет Т., Провост Ф. Data Science для бізнесу. Як збирати, аналізувати і використовувати дані – Київ: Наш Формат, 2019. – 400 с.
2. Маккинни У. Python и анализ данных. Второе издание. – Москва: ДМК Пресс, 2020. - 540с.
3. Вандер П. Дж. Python для сложных задач: наука о данных и машинное обучение. — СПб.: Питер, 2018. — 576 с.: ил.
4. Брюс П., Практическая статистика для специалистов Data Science. 50 важнейших понятий. - СПб.: БХВ-Петербург, 2016. - 304с.
5. Силен Д., Мейсман А., Али М. Основы Data Science і Big Data. Python и наука о данных, СПб.: Питер Пресс, 2019. - 336 с.
6. Джозл Грас, Data Science. Наука про даних з нуля – СПб. : БХВ-Петербург, - 2020. - 416 с.
7. Jake VanderPlas. Python Data Science Handbook: Essential Tools for Working with Data 1st Edition. - O'reilly Media, 2016. - 548 с.

### **Програмне забезпечення навчальної дисципліни**

- Дистрибутив Anaconda та мова програмування Python 3.x.
- MS Excel або інший табличний редактор.
- Дистанційний курс з навчальної дисципліни «Аналіз даних та прикладні пакети статистичної обробки» в системі дистанційного навчання ПУЕТ

### **Політика вивчення навчальної дисципліни та оцінювання**

- Політика щодо термінів виконання та перескладання: завдання, які здаються із порушенням термінів без поважних причин, оцінюються на нижчу оцінку (75 % від можливої максимальної кількості балів за вид діяльності). Перескладання модулів відбувається із дозволу провідного викладача за наявності поважних причин (наприклад, лікарняний).
- Політика щодо академічної доброчесності: списування під час виконання поточних модульних робіт та тестування заборонено (в т. ч. із використанням мобільних девайсів). Мобільні пристрої дозволяється використовувати лише під час он-лайн тестування та підготовки практичних завдань в процесі заняття.
- Політика щодо відвідування: відвідування занять є обов'язковим компонентом. За об'єктивних причин (наприклад, хвороба, працевлаштування, стажування) навчання може відбуватись в он-лайн формі (Moodle) за погодженням із провідним викладачем.

- Політика зарахування результатів неформальної освіти: <http://puet.edu.ua/uk/publiczna-informaciya>

### Оцінювання

Підсумкова оцінка за вивчення навчальної дисципліни розраховується через поточне оцінювання

Вид діяльності	Максимальна кількість балів за вид навчальної роботи
<i>Модуль 1. Основи теорії програмування. Контекстовільні мови. Синтаксичний аналіз.</i>	
<b>Тема 1. Попередній аналіз</b>	
Практичне заняття 1.	6
Практичне заняття 2.	6
Практичне заняття 3.	6
<b>Тема 2. Кореляційний аналіз</b>	
Практичне заняття 4.	6
Практичне заняття 5.	6
<b>Тема 3. Регресійний аналіз</b>	
Практичне заняття 6.	6
Практичне заняття 7.	6
Практичне заняття 8.	6
<b>Тема 4. Дисперсійний аналіз</b>	
Практичне заняття 9.	6
Практичне заняття 10.	6
<b>Тема 5. Дискримінантний аналіз</b>	
Практичне заняття 11.	6
Практичне заняття 12.	6
<b>Тема 6. Кластерний аналіз</b>	
Практичне заняття 13.	11
Модульна контрольна робота №1	
<b>Всього за модулем 1</b>	<b>83</b>
<i>Модуль 2. Аналіз даних з використанням статистичних пакетів</i>	
<b>Тема 7. Аналіз двовимірних даних</b>	
Практичне заняття 14.	6
<b>Тема 8. Класифікація даних</b>	
Практичне заняття 15.	6
Практичне заняття 15. Модульна контрольна робота № 2.	<b>11</b>
<b>Всього за модулем 2</b>	<b>17</b>
<b>Всього по курсу</b>	<b>100</b>

**Шкала оцінювання здобувачів вищої освіти за результатами  
вивчення навчальної дисципліни**

<b>Сума балів за всі види навчальної діяльності</b>	<b>Оцінка за шкалою ЄКТС</b>	<b>Оцінка за національною шкалою</b>
90-100	A	Відмінно
82-89	B	Дуже добре
74-81	C	Добре
64-73	D	Задовільно
60-63	E	Задовільно достатньо
35-59	FX	Незадовільно з можливістю повторного складання
0-34	F	Незадовільно з обов'язковим повторним вивченням навчальної дисципліни