

**Завдання для самостійної роботи з елементами дистанційного навчання  
з дисципліни «Невизначеність та групування інформації»  
на період з 24 січня до 28 лютого 2018 р.**

для студентів

**1 курсу**

**другого (магістерського) рівня**

**освітньої програми «Системи і методи прийняття рішень» (список студентів  
додається)**

**викладач-лектор: д.ф.-м.н., проф. Івохін Є.В. (ел. пошта - [ivohin@univ.kiev.ua](mailto:ivohin@univ.kiev.ua))**

***Види та форми контрольних заходів з перевірки самостійної роботи студентів,  
критерії оцінювання***

Контроль за виконанням самостійної роботи студентами здійснюється у двох формах: у січні-лютому за допомогою електронних засобів (електронною поштою), у березні – шляхом проведення письмової контрольної роботи.

Контроль у січні-лютому 2018 р. відбувається у два етапи. Під час **першого етапу** (24 січня – 6 лютого 2018 р.) студенти мають вивчити питання визначених тем на базовому рівні. Для підтвердження виконання завдання студенти мають надіслати відповіді на контрольні питання викладачу – **Івохіну Є.В.** на електронну пошту [ivohin@univ.kiev.ua](mailto:ivohin@univ.kiev.ua) не пізніше **6 лютого 2018 р.** Викладач оцінює відповіді на контрольні питання в категоріях «зараховано» або «не зараховано». Щоб отримати оцінку «зараховано» потрібно правильно відповісти на всі питання. Якщо студент отримає оцінку «не зараховано», у нього є час до **10 лютого** переробити завдання та надіслати їх викладачу повторно. Також на першому етапі студенти мають обрати собі тему для написання реферату (завдання другого етапу самостійної роботи). Контрольні питання першого етапу, які мають бути виконані та надіслані на електронну пошту викладача, подано **нижче (додаток 1)**.

На **другому етапі** самостійної роботи (7 лютого – 20 лютого 2018 р.) кожен студент має підготувати реферат на одну з наведених тем. Тема реферату погоджується з викладачем.

Реферат є самостійною роботою студента, написаною на основі вивчення наукової літератури та опрацювання джерел. Наукове обґрунтування підходів до розв'язання проблеми є найбільш важливим критерієм високої оцінки реферату. При оцінюванні особлива увага звертається на рівень аналізу проблеми, послідовна аргументація викладеної позиції, переконливі висновки (пропозиції, прогнози). Тексти повністю або частково запозичені з Інтернету чи передруковані з підручника, монографії, статті не зараховуються. Рекомендована тематика рефератів подана в **додатку 2**.

Реферат – форма наукової роботи. За її результатами оформлюється документ, що має стандартну структуру: 1) обсяг 7-10 сторінок друкованого тексту, *інтервал* між рядками 1,5; *шрифт* Times New Roman; *кегель* 14; *вирівнювання тексту* – по ширині; 2) структура: а) назва теми, зміст, ключові слова; б) вступ (актуальність та короткий огляд стану розробленості теми); в) основна частина та короткі висновки; г) перелік найважливішої використаної літератури та джерел.

Реферат має бути надіслане викладачу дисципліни – **Івохіну Є.В.** на електронну пошту [ivohin@univ.kiev.ua](mailto:ivohin@univ.kiev.ua) не пізніше **20 лютого 2018 р.** Викладач оцінює реферат в категоріях «зараховано» або «не зараховано». Викладач повідомляє студенту електронною поштою, чи зараховане його реферат. Якщо реферат не зарахований, викладач вказує недоліки та вимоги щодо його доопрацювання. Не допускається написання реферату на одну й ту ж тему більш ніж двома студентами академічної групи. Ідентичні за змістом реферати отримують оцінку «не зараховано», студенти мають повторно підготувати реферат. У разі незарахування у студента є час до **27 лютого** для його доопрацювання з урахуванням висловлених викладачем зауважень.

Виконання першого етапу самостійної роботи (контрольні питання) є допуском до другого етапу. Виконання другого етапу самостійної роботи (написання реферату) є допуском до написання контрольної роботи у березні 2018 р.

**Якщо відповіді на контрольні питання та реферат здані невчасно без поважних причин або не зараховані, студент втрачає можливість написання контрольної роботи та отримання відповідних модульних балів, без можливості перескладання.**

На контрольну роботу за підсумками самостійної роботи у січні – лютому 2018 р. виносяться всі зазначені вище теоретичні питання. Контрольна робота оцінюється максимум в **10 балів**. Вона включає в себе 2 теоретичних питання з проблематики, винесеної на самостійну роботу. За розгорнуту відповідь на одне теоретичне питання студент може отримати від 1 до 5 балів. Теоретичне питання на контрольній роботі може не співпасти з тим, яке досліджував студент у рефераті.

Критерії оцінювання відповіді студента на одне теоретичне питання:

- повнота розкриття питання 1-2 бали;
- логіка викладення 1 бал;
- використання основної і додаткової літератури 1 бал;
- аналітичні міркування, вміння робити висновки 1 бал.

Контрольна робота буде проведена на першому або другому лекційному занятті з курсу у березні 2018 р.

**Для самостійного опанування студентами у період з 24.01 до 28.02.18 р. виносяться наступні теми та питання, передбачені робочою програмою вибіркової навчальної дисципліни «Невизначеність та групування інформації»:**

#### **Тема 1. Основи кластерного аналізу**

Алгоритми кластеризації, що базуються на розділенні. Ієрархічна кластеризація. Статистична модель кластеризації. Оцінка якості кластеризації.

*Рекомендована література<sup>1</sup>: [основна: 1, 2, 3; додаткова: 1, 2, 3, 4].*

#### **Тема 2. Формалізація даних соціальних мереж**

Схеми здобуття даних з соціальних мереж. Елементи теорії динамічного соціального впливу. Формалізація процесів побудови спільнот. Формалізація даних в задачах кластеризації інформаційних потоків. Кластеризація даних в умовах невизначеності.

*Рекомендована література: [основна: 5; додаткова: 2, 3].*

#### **Тема 3. Методи нечіткої кластеризації**

Метод нечіткої кластеризації C-means. Алгоритм пікового групування нечітких даних. Алгоритм різницевого групування нечітких даних. Метод Кесселя-Густавсона.

*Рекомендована література: [основна: 4; додаткова: 3, 4].*

---

<sup>1</sup> У квадратних дужках вказано порядковий номер у списку основної та додаткової літератури, що наведена в розділі Література та у робочій програмі навчальної дисципліни, яка розміщена на сайті факультету.

**Контрольні питання,**  
**відповіді на які студент надсилає на електронну пошту викладача**  
**([ivohin@univ.kiev.ua](mailto:ivohin@univ.kiev.ua)) не пізніше 6 лютого 2018 р. (перший етап)**  
*(подається коротка, тезисна (до 400 знаків), відповідь на кожне питання).*

**Тема 1. Основи кластерного аналізу**

1. Назвіть основні характеристики алгоритмів кластеризації, що базуються на розділенні.
2. Сформулюйте основні підходи ієрархічної кластеризації.
3. Наведіть основні методи кластеризації.
4. Сформулюйте особливості методики оцінки якості кластеризації.
5. Охарактеризуйте статистичну модель кластеризації.

**Тема 2. Формалізація даних соціальних мереж**

1. Опишіть схеми здобуття даних з соціальних мереж.
2. Сформулюйте основні елементи теорії динамічного соціального впливу.
3. Наведіть схеми процедур побудови спільнот.
4. У чому сутність формалізації в задачах кластеризації інформаційних потоків?
5. Назвіть особливості кластеризації в умовах невизначеності.

**Тема 3. Методи нечіткої кластеризації**

1. Сформулюйте суть методу C-means.
2. Опишіть алгоритм пікового групування.
3. Опишіть алгоритм різницевого групування
4. Основна ідея методу Кесселя-Густавсона.

**Теми для рефератів (другий етап, термін виконання – до 20 лютого 2018 р.)**

1. Кластеризація та класифікація.
2. Ансамблевий підхід до групування даних.
3. Методи ієрархічної кластеризації.
4. Методи кластеризацій нечіткої інформації.
5. Методи кластеризацій складених нечітких даних.
6. Застосування кластеризації для аналізу інформації з соціальних мереж.
7. Задача пошуку спільнот на основі Web-даних.
8. Методи кластеризації графів.
9. Дослідження динаміки процесів кластеризації.
10. Статистичні методи групування даних.

Повний список рекомендованої літератури для написання реферату та підготовки до контрольної роботи можна знайти у робочій програмі з курсу «Невизначеність та групування інформації», розміщеній на сайті факультету комп'ютерних наук та кібернетики.

**Зважаючи на те, що дисципліна вибіркова (з переліків) кожен студент має завчасно пересвідчитись у наявності в списках на вивчення саме цієї дисципліни. Список студентів курсу, які обрали для вивчення дисципліну «Невизначеність та групування інформації» додається.**

**Список студентів 3 курсу, які обрали для вивчення дисципліну  
«Невизначеність та групування інформації»**

1. Галиця Юлія Ігорівна
2. Кулініч Ольга Олегівна
3. Ваврик Петро Русланович
4. Карпенко Владислав Юрійович
5. Лихенко Дар'я Костянтинівна
6. Бесклетна Ольга Сергіївна
7. Буката Андрій Олегович
8. Варивода Андрій Іванович
9. Танасійчук Ростислав Олександрович
10. Філіппов Владислав Михайлович
11. Пришляк Дмитро Олександрович
12. Юркова Катерина Юріївна

**Література**

**Основна**

1. Мандель И. Д. Кластерный анализ. - М.: Финансы и Статистика, 1988. - 616с.
2. Чубукова И.А. Data Mining. Курс лекций интернет-университета INTUIT, 2006. 328 с.
3. Загоруйко Н. Г. Прикладные методы анализа данных и знаний. - Новосибирск: ИМ СО РАН, 1999.
4. Осовский С. Нейронные сети для обработки информации. М: Финансы и статистика, 2002. - 344 с.
5. Ландэ Д.В. Основы интеграции информационных потоков. – К.: Инжиниринг, 2006. - 240 с.

**Додаткова**

1. Классификация и кластер. / Под ред. Дж. Вэн-Райзина. М.: Мир, 1980. 390 с.
2. Kaufman L., Rousseeuw P.J. Finding Groups in Data. An Introduction to Cluster Analysis. Hoboken (NJ): Wiley & Sons Inc., 2005. – 342p.
3. Charu C. Aggarwal, Chandan K. Reddy. Data Clustering: Algorithms and Applications. – CRC Press, 2014.
4. Шлезингер М., Главач В. Десять лекций по статистическому и структурному распознаванию. - Киев: Наукова думка, 2004