

Назва. Інструментальні засоби аналізу даних (мова R)

Тип. Вибіркова.

Рік навчання. 2019–2020.

Семестр. I.

Кількість кредитів ЄКТС. 5.

ПІБ лектора, науковий ступінь, посада. Івахненко О. В., к. е. н., доц.

Результати навчання.

вміння використовувати методи та моделі Data Science для обґрунтування прийняття рішень.

навички роботи у середовищі RStudio та написання програмного коду на мові R.

знання основних методів Data Science та вміння використовувати основні алгоритми маніпулювання даними та побудови моделей складних соціально-економічних процесів та систем.

Обов'язкові попередні навчальні дисципліни. "Теорія ймовірності та математична статистика ", "Статистика", "Економетрика", " прогнозування соціально-економічних процесів".

Зміст.

Вступ до Data Science. Методи збору та вилучення даних. MySQL. Вступ до R. Імпорт, очищення та маніпуляції з даними у R. Візуалізація та розвідковий аналіз у R. Методи навчання без вчителя, з вчителем у R. Машини опорних векторів у R. Побудова дерев рішень. «Випадковий ліс» ("Random forest").

Рекомендовані джерела.

1. Барсегян А. А. Методы и модели анализа данных: OLAP и Data Mining /А. А. Барсегян, М. С. Куприянов, В. В. Степаненко, И. И. Холод. - СПб.: БХВ-Петербург, 2004. - 336 с.
2. Дубров А. М. Многомерные статистические методы / А. М. Дубров, В. С. Мхитарян, Л. И. Трошин. - М.: Финансы и статистика, 1998. - 350 с.
3. Дюк В. Data Mining: учебный курс / Дюк В., Самойленко А. : СПб, 2001. – 368 с.
4. Chambers (2008), Software for Data Analysis, Springer.
5. Venables & Ripley (2002), Modern Applied Statistics with S, Springer.
6. Murrel (2005), R Graphics, Chapman & Hall/CRC Press.

Методи навчання.

Лекції та лабораторні заняття.

Методи оцінювання:

- поточний контроль (комп'ютерне тестування, захист лабораторних робіт);
- модульний контроль (контрольна робота);
- підсумковий контроль (іспит).

Мова навчання. Українська.