



Силабус дисципліни «Методи та моделі Datamining»

Викладач: *Коваленко Ігор Іванович*

професор, д-р. техн. наук, професор кафедри інженерії програмного забезпечення факультету комп'ютерних наук ЧНУ імені Петра Могили.

Очікувані результати навчання

В результаті вивчення дисципліни студент

має знати:

- основні принципи обробки великих масивів даних, способи їх зберігання, візуалізації;
- основні задачі та області застосування інтелектуального аналізу даних;
- основні методи класифікації даних;
- основні методи кластеризації даних;
- основні методи прогнозування даних;
- методи асоціації даних;
- методи візуалізації моделей Data Mining;
- методи аналізу текстових документів Text Mining;
- методи аналізу web-документів web-mining.

має вміти:

- застосовувати моделі, методи та інформаційні технології оперативного й інтелектуального аналізу даних;
- застосовувати методи розробки програмного забезпечення для аналізу, структурування, перетворення та синтезу інформації;
- застосовувати методи Data Mining при розв'язку практичних задач аналізу даних;
- перевіряти побудовані моделі на адекватність, точність отриманих результатів;
- представляти результати аналізу у графічній формі;
- володіти програмними системами інтелектуального аналізу даних.

Обсяг: буде визначено після здійснення процедури вибору студентами вибіркової дисципліни

Мета: засвоєння методів класифікації, кластеризації, прогнозування та асоціації даних; методів візуалізації моделей Data Mining.

Оригінальність навчальної дисципліни: авторський курс

Зміст дисципліни

Тема 1.

Технологія Data Mining.

Тема 2.

Кластерний аналіз даних.

Тема 3.

Основні положення методу міркувань за прецедентами.

Тема 4.

Алгоритми класифікації даних.

Тема 5.

Методи асоціації даних.

Тема 6.

Вилучення знань з web – Web mining.

Тема 7.

Аналіз текстової інформації – Text mining.

Пререквізити

Базується на знаннях дисциплін «Теорія ймовірностей та математична статистика», «Дискретні структури», «Емпіричні методи програмної інженерії», «Чисельні методи», «Алгоритми та структури даних».

Пореквізити

Знання, отримані під час проходження дисципліни, є базою для вивчення таких дисциплін як «Інформаційні технології OLTP, OLAP і Datamining на серверній платформі», «Моделювання та аналіз програмного забезпечення».

Семестровий контроль: залік/іспит/диференційований залік

Оцінювання:

За семестр: 70/60 балів

За залік/іспит: 30/40 балів

Види робіт:

Лабораторні роботи, індивідуальне проєктне завдання

Технічне забезпечення

Проекційне мультимедійне обладнання (проектор, екран, ноутбук/комп'ютер).

Доступ до мережі Internet, точка доступу Wi-Fi.

Програмне забезпечення: MATLAB, RStudio.

Система електронного навчання Moodle 3.9.

Політика щодо дедлайнів

Роботи, які здаються із порушенням термінів без поважних причин, оцінюються на нижчу оцінку.

Політика щодо академічної доброчесності

Передбачає самостійне виконання лабораторних робіт та індивідуального проєктного завдання. Списування під час заліку / іспиту (зокрема із використанням мобільних пристроїв) заборонено. У разі виявлення плагіату або списування роботи не зараховуються.

Критерії оцінювання лабораторних робіт

Вміння застосовувати методи Data Mining при розв'язку практичних задач.

Вміння лаконічно та логічно формувати відповіді на запитання, пов'язані з виконаними роботами, робити запропоновані викладачем зміни в наявному кодї.

Критерії оцінювання самостійної роботи

Повнота та якість виконання завдання.

Якість доповіді студента (форма доповіді, зміст, доказова база, висновки).

Повнота та логічність відповідей на поставлені питання