



СХОВИЩА ДАНИХ ДЛЯ БІЗНЕС-АНАЛІТИКИ

Робоча програма навчальної дисципліни (Силабус)

Реквізити навчальної дисципліни

Рівень вищої освіти	<i>Перший (бакалаврський)</i>
Галузь знань	<i>12 «Інформаційні технології»</i>
Спеціальність	<i>121 «Інженерія програмного забезпечення»</i>
Освітня програма	<i>Інженерія програмного забезпечення інтелектуальних кібер-фізичних систем в енергетиці</i>
Статус дисципліни	<i>Вибіркова</i>
Форма навчання	<i>Очна (денна)</i>
Рік підготовки, семестр	<i>IV курс, осінній семестр</i>
Обсяг дисципліни	<i>120 годин / 3 кредити ECTS</i>
Семестровий контроль/ контрольні заходи	<i>Залік/МКР</i>
Розклад занять	<i>http://rozklad.kpi.ua</i>
Мова викладання	<i>Українська</i>
Інформація про керівника курсу / викладачів	<i>Лектор: к.т.н. Варавя Іван Анрійови, varava_ivan@ill.kpi.ua, @Varavalvan (у робочий час) Практичні заняття: к.т.н. Варавя Іван Анрійови, varava_ivan@ill.kpi.ua, @Varavalvan (у робочий час)</i>
Розміщення курсу	<i>https://campus.kpi.ua</i>

Програма навчальної дисципліни

1. Опис навчальної дисципліни, її мета, предмет вивчення та результати навчання

Силабус навчальної дисципліни «Сховища даних для бізнес-аналітики» складено відповідно до освітньої програми «Інженерія програмного забезпечення інтелектуальних кібер-фізичних систем в енергетиці» підготовки бакалаврів спеціальності 121 – Інженерія програмного забезпечення.

Метою навчальної дисципліни є формування та закріплення у студентів наступних здатностей:

(ЗК 1) Здатність до абстрактного мислення, аналізу та синтезу. (ЗК 2) Здатність застосовувати знання у практичних ситуаціях.

(ФК 3) Здатність розробляти архітектури, модулі та компоненти програмних систем. (ФК 7) Володіння знаннями про інформаційні моделі даних, здатність створювати програмне забезпечення для зберігання, видобування та опрацювання даних.

(ФК 14) Здатність до алгоритмічного та логічного мислення. (ФК 18) Здатність реалізовувати застосунки з використанням концепцій штучного інтелекту, інженерії даних та машинного навчання, зокрема для кібер-фізичних та енергетичних систем.

Предмет навчальної дисципліни – Сервіси MS SQL Server для побудови OLAP-систем на основі сховищ даних.

Програмні результати навчання, на формування та покращення яких спрямована дисципліна: (ПРН 6) Уміння вибирати та використовувати відповідну задачі методологію створення програмного забезпечення; (ПРН 13) Знати і застосовувати методи розробки алгоритмів, конструювання програмного забезпечення та структур даних і знань.

Після засвоєння навчальної дисципліни студенти мають продемонструвати такі результати навчання:

ЗНАННЯ:

- поняття та задачі бізнес-аналітики;
- схеми сховищ даних;
- операції обробки даних у сховищах даних;
- складові процесу інтеграції даних ETL;
- методи інтелектуального аналізу даних у сховищах даних;
- засоби візуалізації інформації із сховищ даних.

УМІННЯ:

- розробляти архітектуру сховищ даних для конкретної предметної області;
- налаштовувати процеси інтеграції даних із різних джерел у сховище даних;
- застосовувати інтелектуальні моделі для аналізу інформації в сховищах даних;
- представляти інформацію із сховищ даних у звітах.

2. Пререквізити та постреквізити дисципліни (місце в структурно-логічній схемі навчання за відповідною освітньою програмою)

Для успішного засвоєння дисципліни студент повинен володіти дисциплінами «Бази даних» та «Компоненти програмної інженерії. Частина 3. Архітектура програмного забезпечення». Компетенції, знання та уміння, одержані в процесі вивчення кредитного модуля «Сховища даних для бізнес-аналітики» можуть бути використані при вивченні дисциплін «Переддипломна практика» та «Дипломне проектування».

3. Зміст навчальної дисципліни

Розділ 1. Теорія сховищ даних

Тема 1.1. Поняття бізнес-аналітики

Тема 1.2. Архітектура сховищ даних

Тема 1.3. Основні елементи сховищ даних

Розділ 2. Багатовимірні моделі даних

Тема 2.1. Багатовимірні куби

Тема 2.2. Мова MDX

Розділ 3. Інтеграція даних у сховища даних

Тема 3.1. ETL-процеси у сховищах даних

Тема 3.2. Застосування сервісу SQL Server Integation Services

Розділ 4. Моделі інтелектуального аналізу даних

Тема 4.1. Огляд сервісу SQL Server Analysis Services

Тема 4.2. Алгоритми Data Mining

Тема 4.3 Інтелектуальні моделі в SSAS

Розділ 5. Звіти в системах бізнес-аналітики

Тема 5.1. Звіти в сервісі Report Services

Тема 5.2. Графічні елементи у звітах Report Services

4. Навчальні матеріали та ресурси

Основна література

1. Alejandro Vaisman, Esteban Zimányi *Data Warehouse Systems: Design and Implementation (Data-Centric Systems and Applications)* / Springer Berlin Heidelberg, 2022. – 696 p.
2. Parteek Bhatia *Data Mining and Data Warehousing: Principles and Practical Techniques* / Cambridge University Press; 1st edition, 2019. – 506 p.
3. Steven Hughes *Hands-On SQL Server 2019 Analysis Services: Design and query tabular and multi-dimensional models using Microsoft's SQL Server Analysis Services* / Packt Publishing. 2020. – 474 p.
4. Christian Cote, Matija Lah, Dejan Sarka *SQL Server 2017 Integration Services Cookbook* / Packt Publishing, 2017. – 558 p.
5. Brian Larson *Microsoft SQL Server 2016 reporting services* / McGraw-Hill Education. 2017. – 849 p.

Додаткова література

6. Jose Chinchilla, Raj Uchanna; Raj Uchhana *Exam Ref 70-767 Implementing a SQL Data Warehouse* Pearson Education Limited (US titles); Microsoft 2018. – 288 p.

Навчальний контент

5. Методика опанування навчальної дисципліни(освітнього компонента)

Лекційні заняття

№ з/п	Назва теми лекції та перелік основних питань (перелік дидактичних засобів, посилання на інформаційні джерела)
1	Тема 1.1. Поняття бізнес-аналітики. Основні питання: Типи користувачів систем бізнес-аналітики. Переваги та недоліки систем бізнес-аналітики. Приклади використання систем бізнес-аналітики.
2	Тема 1.2. Архітектура сховищ даних. Основні питання: Основні вимоги до сховищ даних. Дванадцять правил Кодда для OLAP. Додаткові правила. Особливості(групи правил). Тест FASMI. Моделі OLAP: багатовимірна, реляційна, гібридна. Деякі інші OLAP-моделі.
3	Тема 1.2. Архітектура сховищ даних. Схеми сховищ даних. Основні питання: Підходи Кімбалла та Інмона до побудови сховищ даних. Схема «Зірка». Схема «Сніжинка». Денормалізація бази даних.
4	Тема 1.2. Архітектура сховищ даних. Вітрини даних. Основні питання: Типи вітрин даних. Залежні, незалежні та гібридні вітрини даних. Реалізація вітрин даних. Порівняння сховища даних та вітрини даних
5	Тема 1.3. Основні елементи сховищ даних Основні питання: Вимір. Ієрархічність виміру. Збалансовані, нерівні та незбалансовані ієрархічні виміри.
6	Тема 2.1. Багатовимірні куби Основні питання: Багатовимірна модель даних. Гіперкубічна та полікубічна схеми. OLAP-куб. Розмірність OLAP-кубу. Таблиця фактів. Таблиці вимірів. Операції з кубами: зріз, обертання, консолідація, деталізація..
7	Тема 2.2. Мова MDX Основні питання: Виміри і групи вимірів. Ієрархії та рівні ієрархій. Члени ієрархій. Клітинки. Кортежі. Набори. Осі в MDX.
8	Тема 2.2. Запити в MDX.

	<i>Основні питання: Вибірка в MDX. Вирази MDX. Оператори MDX. Спеціальні оператори MDX. Типи MDX-функцій.</i>
9	Тема 3.1. ETL-процеси у сховищах даних. <i>Основні питання: Поняття ETL-процесу. ETL-процес в системах бізнес-аналітики. Проміжна область в ETL-процесі. Елементи ETL-процесу: витяг, перетворення, завантаження. Очищення даних в ETL.</i>
10	Тема 3.2. Застосування сервісу SQL Server Integration Services. <i>Основні питання: приклад синтезу багатотактної схеми методом таблиць переходів і карт Карно, в якій кількість проміжних змінних більше однієї; правила складання карти відповідності для проміжних змінних.</i>
11	Тема 2.4. Огляд сервісу SQL Server Analysis Services <i>Основні питання: Базова архітектура SQL Server Analysis Services. Об'єкти сервера. Ядро OLAP</i>
12	Тема 2.5. Синтез багатотактних схем методом циклограм. Частина 1 <i>Основні питання: Алгоритми класифікації. Регресивні алгоритми. Алгоритми сегментації. Алгоритми взаємозв'язків. Алгоритми аналізу послідовностей..</i>
13	Тема 2.5. Синтез багатотактних схем методом циклограм. Частина 2 <i>Основні питання: Моделі Data Mining: реляційна та OLAP. Конструктор моделей Data Mining. Аналіз точності: Input Selection, Lift Chart, Classification Matrix, Cross Validation. Прогнозування по моделі.</i>
14	Тема 2.6. Синтез схем з технологічними затримками методом циклограм <i>Основні питання: Основна та вкладені таблиці. Ключовий, вихідний та залежний стовпці. Залежність між атрибутами. Дискретизовані атрибути. Аналіз дерева рішень: вузли та рівні.</i>
15	Тема 2.7. Синтез схем керування на основі тактового розподільника. Частина 1. <i>Основні питання: Кластери. Зв'язок між кластерами. Змінна заливки. Діаграма кластерів. Характеристики кластерів. Порівняння двох кластерів.</i>
16	Тема 2.7. Створення регресійних моделей <i>Основні питання: Модель лінійної регресії. Модель логістичної регресії.</i>
17	Тема 2.8. Звіти в сервісі Report Services <i>Основні питання: Архітектура Report Services. Мова RDL. Утиліта Reporting Server Configuration Manager. Проект звіту в Visual Studio. Налаштування табличного звіту. Засіб перегляду звітів Report Viewer.</i>
18	Тема 2.9. Графічні елементи у звітах Report Services <i>Основні питання: Типи діаграм. Діаграма «Дерево». Діаграма «Сонячні промені».</i> Залік <i>На заліку оголошується кінцева оцінка, яка ставиться у заліково-екзаменаційну відомість. Студенти, що не набрали 60 балів, а також, ті хто хочуть підвищити свою оцінку виконують на занятті залікову контрольну роботу. Студенти, що не допущені до заліку можуть здавати на занятті заборгованості. Якщо недопущений студент зміг протягом заняття отримати допуск та має більш ніж 60 балів, він отримує залікову оцінку на цьому ж занятті. Якщо студент допустився, але 60 балів не набрав, він також має право написати залікову тестову роботу. Недопущені на занятті студенти, а також ті, хто не з'явився на залік і не мають допуску отримують у відомості «не допущений» та відправляються на додаткову сесію. Студенти, що отримали заздалегідь допуск та погоджуються зі своєю оцінкою, можуть не бути присутні на заліковому занятті.</i>

Практичні заняття

№ з/п	Назва теми заняття та перелік основних питань
1	Практичне заняття №1. Створення сховища даних на основі MS SQL Server. Основні питання заняття: Налаштування MS SQL Server для створення сховища даних, проектування схеми сховища даних.
2	Практичне заняття №2. Створення OLAP-кубу для сервісу Analysis Services. Основні питання заняття: уніфікована модель вимірів, виміри кубу, створення проєкцій.
3	Практичне заняття №3. Виконання запитів мовою MDX до сховища даних. Основні питання заняття: MDX-запит, MDX-вираз, оператори мови MDX, функції MDX.
4	Практичне заняття №4. Розробка інтеграційного пакету Integration Services для завантаження даних у сховище даних. Основні питання заняття: Джерела даних, призначення даних, співставлення полів, налаштування SSIS-пакету.
5	Практичне заняття №5. Розробка користувацьких задач обробки даних для ETL-процесу. Основні питання заняття: Задачі Integration Services, клас Task, метод Execute.
6	Практичне заняття №6. Розробка інтелектуальної моделі для кластеризації даних. Основні питання заняття: Алгоритм Microsoft Clustering, перегляд кластерів.
7	Практичне заняття №7. Розробка інтелектуальної моделі для прогнозування даних. Основні питання заняття: Алгоритм лінійної регресії (Microsoft), оцінка точності прогнозу, діаграма точності.
8	Практичне заняття №8. Розробка звіту для сервісу Report Services. Основні питання заняття: Налаштування конфігурації Report Services, розгортання звіту.
9	Практичне заняття №9. Розробка звіту з картографічною інформацією. Основні питання заняття: Map Wizard, MapViewport.

Самостійна робота студента

№ з/п	Вид самостійної роботи	Кількість годин СРС
1	Підготовка до практичних занять	54
2	Підготовка до МКР	6
3	Підготовка до заліку	6

6. Контрольні роботи

Метою контрольних робіт є закріплення та перевірка теоретичних знань із кредитного модуля, набуття студентами практичних навичок самостійного вирішення задач.

Одна модульна контрольна робота (МКР) розбивається на дві контрольні роботи тривалістю в одну годину кожна. Контрольні роботи проводяться за допомогою Google Forms під час календарних контролів (атестацій).

Політика та контроль

6. Політика навчальної дисципліни (освітнього компонента)

Система вимог, які викладач ставить перед студентом:

- правила відвідування занять: відповідно до Наказу 1-273 від 14.09.2020 р. заборонено оцінювати присутність або відсутність здобувача на аудиторному занятті, в тому числі нараховувати заохочувальні або штрафні бали. Відповідно до РСО даної дисципліни бали нараховують за відповідні види навчальної активності на лекційних та практичних заняттях.

- правила поведінки на заняттях: студент має можливість отримувати бали за відповідні види навчальної активності на лекційних та практичних заняттях, передбачені РСО дисципліни. Використання засобів зв'язку для пошуку інформації на гугл-диску викладача, в інтернеті, в дистанційному курсі на платформі Сікорський здійснюється за умови вказівки викладача;

- політика дедлайнів та перескладань: якщо студент не проходив або не з'явиться на МКР (без поважної причини), його результат оцінюється у 0 балів. Перескладання результатів МКР не передбачено;

- політика щодо академічної доброчесності: Кодекс честі Національного технічного університету України «Київський політехнічний інститут» <https://kpi.ua/files/honorcode.pdf> встановлює загальні моральні принципи, правила етичної поведінки осіб та передбачає політику академічної доброчесності для осіб, що працюють і навчаються в університеті, якими вони мають керуватись у своїй діяльності, в тому числі при вивченні та складанні контрольних заходів з дисципліни «Сховища даних для бізнес-аналітики»;

- при використанні цифрових засобів зв'язку з викладачем (мобільний зв'язок, електронна пошта, переписка на форумах та у соцмережах тощо) необхідно дотримуватись загальноприйнятих етичних норм, зокрема бути ввічливим та обмежувати спілкування робочим часом викладача.

7. Види контролю та рейтингова система оцінювання результатів навчання (РСО)

Поточний контроль: тестування, МКР, виконання завдань до практичних занять.

Календарний контроль: провадиться двічі на семестр як моніторинг поточного стану виконання вимог силабусу.

Семестровий контроль: залік.

Умови допуску до семестрового контролю: семестровий рейтинг більше 40 балів.

Таблиця відповідності рейтингових балів оцінкам за університетською шкалою:

Кількість балів	Оцінка
95-100	Відмінно
85-94	Дуже добре
75-84	Добре
65-74	Задовільно
60-64	Достатньо
Менше 60	Незадовільно
Менше 30	Не допущено

Загальна рейтингова оцінка студента після завершення семестру складається з балів, отриманих за:

- тестування по кожному лекційному заняттю;
- виконання завдань до практичних занять;
- виконання модульних контрольних робіт (МКР);
- виконання додаткових завдань.

Тестування по лекціям	Практичні заняття	МКР 1	МКР 2	Додаткові бали
18	54	10	10	8

Тестування по матеріалам лекційних занять

Ваговий бал 1. Максимальна кількість балів за тестування – 1 бал * 18 лекцій = 18 балів.

Тестування проводиться за допомогою Google Forms наприкінці поточної лекції. У деяких випадках термін проходження тестування може бути продовжений лектором. Тривалість проходження одного тестування – 5 хвилин. Кількість спроб – одна. У деяких випадках, що пов'язані з технічними проблемами студентів, може надатися повторна спроба на окремі тестування.

Кожне тестування містить 5 запитань різного формату (вибір правильного варіанту з переліку; вибір правильних варіантів з переліку, вірно/невірно; визначити відповідність; чисельна відповідь).

Критерії оцінювання

- запитання типу «вибір правильного варіанту з переліку», «вірно/невірно», «чисельна відповідь» оцінюються однозначно: вірна відповідь – 0,2 бал, невірна відповідь – 0 балів;
- запитання, на які немає однієї конкретної відповіді, типу «вибір правильних варіантів з переліку» або «визначити відповідність» оцінюються у відповідності до кількості елементів у тесті (наприклад, якщо потрібно обрати 4 вірні варіанти відповіді із 7, то студент отримає по 0,05 балів за один правильний варіант відповіді, а за всі 4 правильні відповіді отримає відповідно 0,2 бал) – невірна відповідь – 0 балів, частково вірна відповідь – 0,05-0,15 балів, вірна відповідь 0,2 бал.

Практичні заняття

Ваговий бал – 6. Максимальна кількість балів за всі практичні заняття – 6 бали * 9 занять = 54 балів.

На практичних заняттях студенти розробляють код мовою VBA у відповідності до тематики практичного заняття. Після кожного практичного заняття студенти отримують домашнє завдання, яке необхідно вирішити та надати на перевірку викладачу до початку наступного заняття (зазвичай це 2 тижні, однак іноді цей час може бути змінений викладачем у деяких конкретних випадках).

Критерії оцінювання

- домашнє завдання вирішено вірно та здано протягом 2-х тижнів після практичного заняття – 6 балів;
- домашнє завдання вирішено вірно, але здано протягом більш ніж 2-х тижнів після практичного заняття – 5 балів;
- домашнє завдання вирішено із незначними помилками та здано протягом 2-х тижнів після практичного заняття – 4 бали;

- домашнє завдання вирішено із незначними помилками та здано протягом більш ніж 2-х тижнів після практичного заняття – 3 бали;
- домашнє завдання вирішено із значними помилками – повертається на доопрацювання.

Модульна контрольна робота

Ваговий бал за одну МКР – 10. Максимальний бал за 2 МКР складає 20 балів.

Критерії оцінювання

Кожна частина модульної контрольної роботи складається із 20 завдань. Окреме завдання оцінюється в 0.5 бали.

Календарний контроль базується на поточній рейтинговій оцінці. Умовою позитивної атестації є значення поточного рейтингу студента не менше 50% від максимально можливого на час атестації.

Додаткові (бонусні) бали

Рейтинговою системою оцінювання передбачені додаткові бали за виконання додаткових завдань. Один студент не може отримати більше ніж 8 бонусних балів у семестрі. При отриманні більш ніж 8 балів, вони обмежуються на рівні 8. Бонусні бали можуть бути отримані за виконання «Завдання до лекцій».

Завдання до лекцій. Студенти, за бажанням, можуть виконувати додаткові завдання за матеріалами лекцій (розробити макрос чи користувацьку функцію). За одне додаткове завдання нараховується 0,5 бали. Максимальна кількість балів, що можна отримати за завдання до лекцій складає 8 балів.

Форма семестрового контролю – залік

Максимальна сума балів складає 100. Необхідною умовою допуску до заліку є зараховані обидві частини МКР та здані усі завдання до практичних занять. Для отримання заліку з кредитного модулю «автоматом» потрібно мати рейтинг не менше 60 балів, а також виконані умови допуску до заліку.

Студенти, які наприкінці семестру мають рейтинг менше 60 балів, а також ті, хто хоче підвищити свою оцінку в системі ECTS, виконують залікову контрольну роботу. При цьому набрані бали студентом анулюються, а оцінка за залікову контрольну роботу є остаточною.

Залікова робота. Залікова робота проводиться на останньому лекційному занятті. Студент проходить тестування за допомогою Google Forms. На тестування пропонується 50 тестових завдань, кожне з яких оцінюється в 2 бали. Для отримання позитивної оцінки необхідно набрати 60 балів і вище. Час тестування зазвичай складає 90 хвилин, але може бути скоригований лектором та (або) викладачам, що приймає залік.

Робочу програму навчальної дисципліни (силабус):

Складено доцентом кафедри інженерії програмного забезпечення в енергетиці, к.т.н. Варавою І.А.

Ухвалено кафедрою інженерії програмного забезпечення в енергетиці (протокол № 28 від 15.05.2023 р.)

Погоджено Методичною комісією ННІАТЕ КПІ ім. Ігоря Сікорського (протокол № 9 від 26.05.2023 р.)