

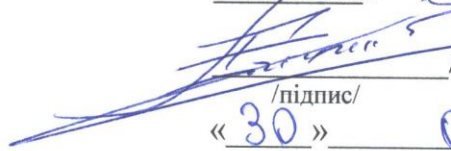
МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ «ЛЬВІВСЬКА ПОЛІТЕХНІКА»

Інститут комп'ютерних наук та інформаційних технологій
/назва навчально-наукового інституту/

Кафедра інформаційних систем та мереж
/назва /

«ЗАТВЕРДЖУЮ»

Голова науково-методичної комісії
спеціальності «Інформаційні системи та
технології»

 / О.В. Стетук /
/підпис/ /ініціали та прізвище /
« 30 » 08 20 19 року

РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

СК7 «Методи опрацювання великих даних»

/код і назва навчальної дисципліни/

Другий (магістерський)

/рівень вищої освіти/

галузь знань 12 Інформаційні технології

/шифр і назва/

спеціальність 126 «Інформаційні системи та технології»

/шифр і назва /

спеціалізація _____

/шифр і назва /

вид дисципліни обов'язкова

(обов'язкова / за вибором)

мова викладання українська

Львів – 2019 рік

Робоча програма з навчальної дисципліни «Методи опрацювання великих даних» для студентів
Інституту комп'ютерних наук та інформаційних технологій
/назва інституту/

Розробники:

Професор кафедри ІСМ, д.т.н, професор  / А. Ю. Берко /
/посада, науковий ступінь та вчене звання/ /підпис/ /ініціали та прізвище/

Робочу програму розглянуто та схвалено на засіданні кафедри ІСМ
Протокол від «28» 08 2019 року № 1

Завідувач кафедри ІСМ  / В.В.Литвин /
/підпис/ /ініціали та прізвище /

Робочу програму розглянуто та схвалено НМК спеціальності
126 «Інформаційні системи та технології»

Протокол від «26» 08 2019 року № 1

Секретар НМК  / Т.В.Местаків /
/підпис/ /ініціали та прізвище /

1. Структура навчальної дисципліни

Найменування показників	Всього годин	
	Денна форма навчання	Заочна форма навчання
Кількість кредитів/год.	4/120	4/120
Усього годин аудиторної роботи, у т.ч.:	48	20
• лекційні заняття, год.	16	6
• семінарські заняття, год.	—	
• практичні заняття, год.	—	
• лабораторні заняття, год.	32	14
Усього годин самостійної роботи, у т.ч.:	72	100
• курсова робота.	-	-
• розрахункові (розрахунково-графічні) роботи, к-сть/год.	—	-
• підготовка до навчальних занять та контрольних заходів	32	50
• Підготовка до лабораторних робіт	40	50
Екзамен	+	+
Залік		

Частка аудиторного навчального часу студента у відсотковому вимірі:

- денної форми навчання – 40%;
- заочної форми навчання – 16,7%

2. Мета та завдання навчальної дисципліни

2.1. Мета вивчення навчальної дисципліни

Мета дисципліни – вивчення системних методів формування вимог та управління ними в процесі виконання проекту, методів та засобів моделювання вимог

Набуття поглиблених теоретичних та фундаментальних знань для ефективного розв'язування задач опрацювання великих даних під час професійної діяльності у різних галузях або у процесі навчання на ґрунті коректного формулювання задач та аналізу великих даних практичних розв'язання практичних проблем, що передбачає застосування теорій та методів опрацювання великих даних, а також комп'ютерного моделювання інформаційних систем.

Надати студентам теоретичні знання та практичні навички оволодіння та використання великих даних, та комп'ютерних технологій отримання та аналізу великих даних, створення і впровадження інформаційних систем нового покоління на основі великих даних

Забезпечити студентам здатність виконувати аналіз великих даних будь-якого масштабу в умовах високої невизначеності, що викликається запитом на зміни і ризиками, з урахуванням впливу зовнішнього оточення; розроблення нових інструментів і методів аналізу великих даних

2.2. Завдання навчальної дисципліни

Внаслідок вивчення навчальної дисципліни студент повинен бути здатним продемонструвати такі **результати навчання**:

1. володіння знаннями інформаційних систем і технологій в аналізі та прийнятті рішень;
2. вміння реалізувати загальний процес опрацювання великих даних, будувати системні моделі великих даних;
3. використовувати структурний та об'єктний підходи та формальні методи в аналізі великих даних, визначати розширені зв'язки та закономірності в даних;
4. здатність формувати склад і зміст інформаційних ресурсів великих даних для ефективного аналізу даних в різних предметних областях;
5. здатність формувати теоретичні та практичні рішення в опрацюванні великих даних

6. здатність використовувати знання та навички щодо проведення збору даних, моделювання відповідних ресурсів і систем при аналізі великих даних
7. володіння методами опрацювання та аналізу великих даних в різних галузях;
8. здатність застосувати знання та практичні навички аналізу відповідних нормативних документів, чинних стандартів і технічних умов у галузі застосування великих даних;
9. вміння застосовувати сучасні знання та новітні технології в галузі аналізу великих даних;

Вивчення навчальної дисципліни передбачає формування та розвиток у студентів компетентностей:

загальних:

ІНТ Здатність використовувати поглиблені теоретичні та фундаментальні знання, уміння і навички для успішного розв'язування складних спеціалізованих задач та практичних проблем під час професійної діяльності у галузі інформаційних систем та технологій або у процесі навчання, що передбачає застосування певних теорій та методів відповідної науки і характеризується комплексністю та невизначеністю умов.

ЗК1 Здатність до абстрактного мислення, аналізу та синтезу на відповідних рівнях.

ЗК2 Здатність застосовувати знання у практичних ситуаціях.

ЗК4 Здатність спілкуватися, читати та писати іноземною мовою.

ЗК5 Навички використання інформаційних і комунікаційних технологій.

ЗК6 Здатність вчитися і оволодівати сучасними знаннями.

ЗК7 Здатність до пошуку, оброблення та узагальнення інформації з різних джерел.

ЗК8 Здатність працювати в команді та особисто.

ЗК9 Навички міжособистісної взаємодії.

ЗК10 Здатність розробляти та управляти проектами.

ЗК11 Навички здійснення безпечної діяльності.

ЗК12 Здатність оцінювати та забезпечувати якість виконуваних робіт.

фахових:

ФК1 Здатність проводити аналіз об'єкту проектування та предметної області.

ФК2 Володіння навчально-методичними основами і стандартами в області ІСТ, уміння їх застосовувати при розробці функціональних профілів ІСТ, при побудові та інтеграції систем, продуктів і сервісів ІСТ.

ФК3 Здатність до проектування системного, комунікаційного і прикладного програмного забезпечення, технічних засобів та комунікаційних й інформаційних технологій, мереж та систем.

ФК4 Здатність розробляти засоби реалізації ІСТ (методичні, інформаційні, алгоритмічні, технічні й програмні).

ФК6 Здатність використовувати сучасні технології проектування в розробці алгоритмічного та програмного забезпечення ІСТ.

ФК7 Здатність застосовувати, впроваджувати та експлуатувати сучасні ІСТ (виробничі, підтримки прийняття рішень, інтелектуального аналізу даних) у різних галузях людської діяльності, національної економіки та виробництва.

ФК9 Здатність управляти якістю продуктів і сервісів ІСТ протягом їх життєвого циклу.

ФК10 Здатність проводити оцінку виробничих і невиробничих витрат на забезпечення якості об'єкта проектування, розробляти бізнес-рішення та оцінювати нові технологічні пропозиції.

ФК12 Здатність здійснювати організацію робочих місць, їх технічне оснащення, розміщення комп'ютерного устаткування, використання організаційних, технічних, алгоритмічних та інших методів організаційно-управлінської діяльності.

ФК13 Здатність формулювати та коректно ставити завдання та керувати молодшим технічним персоналом; пов'язувати технічні та управлінські підрозділи організації, а також брати активну участь у навчанні користувачів.

ФК14 Здатність розуміти, розгортати, організовувати, управляти та користуватися сучасними навчально-дослідницькими ІСТ (у тому числі такими, що базуються на використанні Інтернету), інформаційними та комунікаційними технологіями.

Результати навчання даної дисципліни деталізують такі **програмні результати навчання:**

ЗН1 Здатність використовувати поглиблені професійно-профільні знання та практичні навичками для оптимізації проектування інформаційних систем будь-якої складності, для вирішення конкретних завдань проектування інтелектуальних інформаційних систем з управління об'єктами різної фізичної природи.

ЗН2 Здатність формулювати та вдосконалювати важливу дослідницьку задачу, для її вирішення збирати необхідну інформацію та формулювати висновки, які можна захищати в науковому контексті.

ЗН3 Здатність проводити системний аналіз об'єктів проектування та обґрунтовувати вибір структури, алгоритмів та способів опрацювання інформації в ІСТ.

ЗН6 Здатність брати участь у проектуванні ІСТ, мати базові знання зі змісту і правил оформлення проектних матеріалів, знати склад та послідовність виконання проектних робіт з урахуванням вимог відповідних нормативно-правових документів.

ЗН7 Вміти обґрунтовувати вибір технічної структури та розробляти відповідне програмне забезпечення, що входить до складу ІСТ.

ЗН9 Здатність розуміти і враховувати соціальні, екологічні, етичні, економічні аспекти, вимоги охорони праці, виробничої санітарії, пожежної безпеки та наявних державних і закордонних стандартів під час формування технічних завдань та рішень.

ЗН10 Здатність демонструвати вміння розробляти техніко-економічне обґрунтування розроблення ІСТ та вміти оцінювати економічну ефективність їх впровадження.

УМ1 Уміння спілкуватись англійською мовою в обсязі, достатньому для здійснення професійної діяльності, читання та трактування міжнародних технічних стандартів.

УМ2 Здатність до аналізу предметної області та синтезу інформаційних систем та технологій із використанням сучасних методів та засобів інформаційних технологій.

УМ3 Вміти застосовувати методи пошуку джерел інформації; аналізувати якість отриманої інформації.

КОМ1 Уміння спілкуватись, включаючи усну та письмову комунікацію українською та іноземною мовами (англійською, німецькою, італійською, французькою, іспанською).

КОМ2 Здатність використання різноманітних методів, зокрема сучасних інформаційних технологій, для ефективно спілкування на професійному та соціальному рівнях.

АіВ1 Здатність адаптуватись до нових ситуацій та приймати відповідні рішення.

АіВ2 Здатність усвідомлювати необхідність навчання впродовж усього життя з метою поглиблення набутих та здобуття нових фахових знань.

АіВ3 Здатність відповідально ставитись до виконуваної роботи, самостійно приймати рішення, досягати поставленої мети з дотриманням вимог професійної етики.

2.3. Перелік попередніх та супутніх і наступних навчальних дисциплін

№ з/п	Попередні навчальні дисципліни	Супутні і наступні навчальні дисципліни
1.	Організація баз даних	Інноваційні інформаційні технології
2.	Системи управління базами даних	Технології інтеграції інформаційних ресурсів

3. Анотація навчальної дисципліни

У викладанні дисципліни розглядаються такі теми.

Поняття та визначення Великих даних. Розвиток концепції Великих даних. Методики аналізу Великих даних. Технології аналізу Великих даних. Моделі зберігання Великих даних. Модель обчислень MapReduce. Засоби опрацювання Великих даних. Застосування Великих даних у предметних областях.

4. Опис навчальної дисципліни

4.1. Лекційні заняття

№ п/п	Найменування розділів, тем	Кількість годин	
		ДНФ	ЗНФ
1.	Концепція Великих даних	2	2
	Поняття та визначення Великих даних. Властивості Великих даних. Вимоги до Великих даних. Специфіка Великих даних.		
2.	Розвиток концепції Великих даних.	2	
	Опрацювання великих обсягів даних на MainFrame. Передумови та чинники виникнення напряду Великих даних. Становлення та розвиток технологій Великих даних. Предметні області застосування великих даних. Сучасний стан та перспективи розвитку Великих даних.		2
3.	Методики аналізу Великих даних.	2	
	А/В тестування. Класифікація. Кластерний аналіз. Краудсорсинг (відбір даних). Зміщування та інтеграція даних. Data mining. Визначення узгоджень(гармонійності) даних. Генетичні алгоритми. Машинне навчання. Опрацювання природної мови. Мережевий аналіз. Оптимізація. Розпізнавання шаблонів. Прогнозне моделювання. Регресійний аналіз. Опрацювання сигналів. Просторовий аналіз даних. Статистика. Імітаційне моделювання (Симуляція). Аналіз часових послідовностей. Вивчення асоціативних в'язків. Вивчення функціональних зав'язків. Вивчення прихованих в'язків.		
4.	Технології аналізу Великих даних	2	2
	Мова R; 1010data; Apache Chukwa; Apache Hadoop; Apache Hive; Apache Pig!; Jaspersoft; LexisNexis Risk Solutions HPCC Systems; MapReduce; Revolution Analytics (на базі мови R для мат.статистики).		
5.	Моделі зберігання Великих даних	2	
	Потокові дані. Big Table Великі дані. HBase-сховища. Cassandra/ Спеціалізовані сховища Великих даних. SQL для Великих даних.		2
6.	Модель обчислень MapReduce	2	
	Парадигма MapReduce. Модель обчислень MapReduce. Фази Map, Shuffle, Reduce.		
7.	Засоби опрацювання Великих даних	2	
	Apache Hadoop.		
8.	Застосування Великих даних у предметних областях.	2	
	Екологічний моніторинг. Соціальні процеси. Державне управління. Маркетинг. Торгівля. Е-комерція. Медицина. Біржова діяльність. Політика.		6
	ЗАГАЛОМ	16	

4.2. Лабораторні заняття

№ п/ п	Зміст занять	Кількість годин	
		ДНФ	ЗНФ
1.	Інтерфейс користувача Apache Hadoop	4	1
2.	Налаштування параметрів середовища Apache Hadoop	4	1
3.	Конфігурація завдань Apache Hadoop	4	2
4.	Виконання задач в Apache Hadoop	4	2
5.	Моніторинг завдань в Apache Hadoop	4	2
6.	Ввід даних в Apache Hadoop	4	2
7.	Вивід даних в Apache Hadoop	4	2
8.	Документування результатів в Apache Hadoop	4	2
	ЗАГАЛОМ	32	14

4.3. Самостійна робота

№ з/п	Найменування робіт	Кількість годин	
		ДФН	ЗФН
1.	підготовка до навчальних занять та контрольних заходів	32	50
2.	підготовка до лабораторних робіт	40	50
Загалом годин		72	100

5. Методи діагностики знань

Діагностика знань відбувається шляхом оцінювання виконаних лабораторних робіт та іспиту (письмова компонента – тестові запитання трьох рівнів складності та усна компонента – опитування).

6. Критерії оцінювання результатів навчання студентів

Розподіл балів у 100-бальній шкалі		
Поточний контроль (ПК)	Іспит	Разом за Дисципліну
Виконання лабораторних робіт		
40	60	100

7. Навчально-методичне забезпечення

Конспект лекцій, методичні вказівки до виконання лабораторних робіт.

8. Рекомендована література

Базова

1. White, Tom // Hadoop: The Definitive Guide // O'Reilly Media, 2009.
2. Hadoop. Apache Software Foundation // <http://hadoop.apache.org/>
3. Finley, Klint // Steve Ballmer on Microsoft's Big Data Future and More in This Week's Business Intelligence Roundup // ReadWriteWeb, 2011.
4. Fay Chang, Jeffrey Dean, Sanjay Ghemawat & etc. // Bigtable: A Distributed Storage System for Structured Data // Google Lab, 2006.
5. Сухорослов, О. // Новые технологии распределенного хранения и обработки больших массивов данных // Институт системного анализа РАН, 2008.
6. Jeffrey Dean, Sanjay Ghemawat // MapReduce: Simplified Data Processing on Large Clusters // Google Inc., 2004.

7. Judy Qiu // Cloud Technologies and Their Applications // Indiana University Bloomington, 2010
8. The Hadoop Distributed File System: Architecture and Design // http://hadoop.apache.org/common/docs/r0.17.2/hdfs_design.html
9. Созыкин, А. // Параллельное программирование в Hadoop // <http://www.asozykin.ru/courses/hadoop>
10. Ralf Lammel // Google's MapReduce Programming Model — Revisited // Microsoft Corp.

9. Допоміжна

1. Douglas, L. // 3D Data Management: Controlling Data Volume, Velocity and Variety // Gartner, 2001.
2. Christy Pettey, Laurence Goasduff // Gartner Says Solving 'Big Data' Challenge Involves More Than Just Managing Volumes of Data // Gartner, 2011.
3. Дубова, Н. // Большая конференция о Больших Данных // Открытые системы, 2011.
4. Henschen, Doug // Oracle Releases NoSQL Database, Advances Big Data Plans // InformationWeek, 2011.
5. Шах, Агам // HP меняет персональные компьютеры на Большие Данные // Открытые системы, 2011.
6. Doug Henschen // EMC Tries To Unify Big Data Analytics // InformationWeek, 2011.

9. Інформаційні ресурси

1. MapReduce. Wikipedia, the free encyclopedia // <http://ru.wikipedia.org/wiki/MapReduce>
2. Новікова М. Інформаційні технології: історія та сучасність [Електронний ресурс] / Марія Новікова // Рівненська обласна універсальна наукова бібліотека: сайт. – Режим доступу : <http://libr.rv.ua/ua/virt/133/>.
3. Плєскач В.Л., Затонацька Т.Г. Інформаційні системи і технології на підприємствах [Електронний ресурс] – Плєскач В.Л. . – Режим доступу: http://pidruchniki.com/1059110247701/informatika/informatsiyi_sistemi_i_tehnologiyi_na_pidpriyemstva
4. ISO [Електронний ресурс] // Офіційний сайт ISO. – Режим доступу : <http://www.iso.org/iso/home.htm>.

10. Узгодження з іншими навчальними дисциплінами

№ з/п	Назва навчальної дисципліни, щодо якої проводиться узгодження	Прізвище та ініціали викладача	Підпис
1.	Техн. проєкт. сист. бізнес-аналіст	Кіс Є.П.	Откіс
2.	Безпека розпод. інформ. систем	Кравець П.О.	Кравець

11. Зміни та доповнення до робочої програми навчальної дисципліни

№ з/п	Зміст внесених змін (доповнень)	Дата і № протоколу засідання кафедри	Примітки
1			
2			
3			