


МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
 Чорноморський національний університет імені Петра Могили
 Факультет комп'ютерних наук
 Кафедра інтелектуальних інформаційних систем

“ЗАТВЕРДЖУЮ”

Перший проректор
 Іщенко Н.М.

“28”  2020 року

**РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ
 «МЕТОДИ ЗБОРУ ТА ОБРОБКИ ДАНИХ З МЕРЕЖІ INTERNET»**

Спеціальність 076 «Підприємництво, торгівля та біржова діяльність»

Освітня програма «Економіка та управління підприємством»

Рівень вищої освіти – другий (магістерський)

Розробник
 Завідувач кафедри інтелектуальних
 інформаційних систем
 Завідувач кафедри економіки та підприємництво
 Гарант освітньої програми
 В.о. декана факультету економічних наук
 Директор ННІПО
 Начальник НМВ

Кулаковська І.В.
 Кондратенко Ю.П.
 Кузьменко О.Б.
 Кузьменко О.Б.
 Філімонова О.Б.
 Норд Г.Л.
 Шкірчак С.І.



Миколаїв – 2020 рік

1. Опис навчальної дисципліни

Найменування показника	Характеристика дисципліни	
Найменування дисципліни	Методи збору та обробки даних з мережі Internet	
Галузь знань	07 «Управління та адміністрування»	
Спеціальність	076 «Підприємництво, торгівля та біржова діяльність»	
Спеціалізація (якщо є)	-	
Освітня програма	«Економіка та управління підприємством»	
Рівень вищої освіти	Другий (магістерський)	
Статус дисципліни	Вибіркова	
Курс навчання	V	
Навчальний рік	2020-2021	
Номер семестрів:	Денна форма	Заочна форма
	11 семестр	11
Загальна кількість кредитів ЄКТС/годин	3 кредитів / 90 годин	
Структура курсу:	Денна форма	Заочна форма
	– лекції	6
	– практичні	8
	– годин самостійної роботи студентів	76
Відсоток аудиторного навантаження	33%	16%
Мова викладання	Українська	
Форма проміжного контролю	Тестування	
Форма підсумкового контролю	Залік	

2. Мета, завдання та результати вивчення дисципліни

Концепція викладання дисципліни:

Дослідження методів збору та обробки даних з мережі INTERNET з кожним роком набуває все більшої актуальності у зв'язку з необхідністю забезпечення безпеки населення і моніторингу громадських настроїв. При цьому слід зазначити важливість аналізу індивідуального масиву даних кожного користувача, що включає інформацію про його взаємодію, про контент, сформованим цим користувачем, а також інформацію про взаємодію з контентом інших користувачів. На основі результатів аналізу цього масиву даних можна зробити висновки про особу користувача, його характері і звичаї, що сприяє оцінці ситуації при самоорганізації суспільства і його окремих груп.

Концепція дисципліни «МЕТОДИ ЗБОРУ ТА ОБРОБКИ ДАНИХ З МЕРЕЖІ INTERNET» базується на методах одномірного та багатомірного аналізу даних та особливостей їх використання, які використовуються для моделювання матеріальних, фінансових, інформаційних потоків. При цьому вивчення поєднується із використанням сучасних інформаційних технологій для вирішення задач візуалізації даних та огляду методів та інструментів для аналізу: соціальних графів, тональності тексту, аудіо та відео об'єктів.

«МЕТОДИ ЗБОРУ ТА ОБРОБКИ ДАНИХ З МЕРЕЖІ INTERNET» є вибірковою для циклу професійної підготовки фахівців освітньо-кваліфікаційного рівня «магістр» за спеціальністю 076 Підприємництво, торгівля та біржова діяльність

Вивчаються: методи збору даних з мережі INTERNET, соціальних мереж, методи обробки даних, методи визначення соціальних груп на основі даних із соціальних мереж, класифікація вершин графу. Знання та вміння, отримане при вивченні цієї дисципліни допомагають студентам розбиратися в сучасних методах та алгоритмах, які використовуються при зборі, аналізі та обробці даних з INTERNET, соціальних мереж, зберігати їх в графах, попередньо типізувати дані та приводити їх до потрібного вигляду, працювати з бібліотеками аналізу даних, вміти працювати з інструментами для візуалізації даних.

Мета вивчення дисципліни – формування у студентів знань та вмінь використання сучасних технологій аналізу даних у та вдосконалення у студентів теоретичних знань та практичних навиків, щодо прийняття рішень в області управління інформаційними потоками і технологіями аналізу даних для вирішення різноманітних задач.

Задача дисципліни – необхідно сформулювати у студентів відповідні вміння і навички з статистичних методів обробки даних з INTERNET, їх досліджень, аналізу документів, а також сформулювати навички роботи на основі конкретних емпіричних даних.

Завдання дисципліни – отримання студентами теоретичних знань та практичних навичок щодо застосування методів та алгоритмів в області управління інформаційними потоками і технологіями аналізу даних для вирішення різноманітних задач:

- виконувати всі процедури проведення дослідження: починаючи з складання програми дослідження, виділення категорій аналізу, збір інформації, обробку і інтерпретацію результатів дослідження;
- знати та уміти застосовувати методи аналізу інформації, їх оцінка та використання при аналізі ризиків поширення фейкової інформації в мережах.
- збір даних з соціальних мереж, методологію, методи, техніку та технологію проведення та організації дослідження за допомогою інструментів аналізу даних та бібліотек сучасних мов програмування;
- вміти працювати з інструментами для візуалізації даних.

Освоєння даної дисципліни направлене на формування у студента наступних загальних та спеціальних (фахових, предметних) компетентностей.

Загальні компетентності:

ЗК 2. Вміння виявляти, ставити та вирішувати проблеми.

Спеціальні (фахові, предметні) компетентності:

СК 2. Здатність проводити оцінювання продукції, товарів і послуг в підприємницькій, торговельній та/або біржовій діяльності.

СК 4. Здатність до вирішення проблемних питань і прийняття управлінських рішень у професійній діяльності.

Освоєння даної дисципліни направлене на формування у студента наступних **програмних результатів навчання:**

РН 2. Визначати, аналізувати проблеми підприємництва, торгівлі і біржової діяльності та розробляти заходи щодо їх вирішення.

РН 7. Оцінювати продукцію, товари, послуги, а також процеси, що відбуваються в підприємницьких, торговельних та/або біржових структурах, і робити відповідні висновки для прийняття управлінських рішень.

РН 9. Вміти вирішувати проблемні питання, що виникають в діяльності підприємницьких, торговельних та/або біржових структур за умов невизначеності та ризиків.

3. Програма навчальної дисципліни

Денна форма:

№	Теми	Лекції	Практичні	Самостійна робота
1	Масові системи аналізу даних з INTERNET	2	4	12
2	Системи моніторингу і аналізу даних з INTERNET для комерційних організацій	2	4	12
3	Візуальне представлення структур соціальних мереж. Засоби автоматичного розміщення при візуалізації графів	2	4	12
4	Знаходження релевантної інформації. Виведення знань з інформації, яка міститься у INTERNET.	2	4	12
5	Персоналізація інформації. Знання про споживачів / Користувачів.	2	4	12
	Всього за курсом	10	20	60

4. Зміст навчальної дисципліни

4.1. План лекцій

№	Тема заняття / план
1	<p>Лекція 1. Масові системи аналізу даних з INTERNET</p> <p>1. Пошукові системи: пошук в Твіттері: search.twitter.com; пошук в блогах; пошук людей в соціальних мережах.</p> <p>2. Інформаційні системи, що використовують пошукові машини для пошуку документів за запитом користувача і регулярно відсилають найбільш релевантні результати пошуку на електронну пошту користувача: GoogleAlerts (працює поверх пошукової машини Google), Twilert (працює поверх пошукової машини Twitter).</p> <p>3. Системи збору інформації по заданій множині джерел (RSS-підписка): GoogleReader. Системи збору і об'єднання інформації з різних джерел («мішані» дані):</p>
2	<p>Лекція 2. Системи моніторингу і аналізу даних з INTERNET для комерційних організацій</p> <p>1. Системи моніторингу і аналізу всередині організації (соціальна мережа організації або вже явно визначена, або ідентифікується якимось чином)</p> <p>2. Системи моніторингу і аналізу за межами організації: дослідження ринку (аналіз присутності в соціальних мережах брендів і їх рішень, аналіз думок користувачів про бренди і рішеннях, визначення характеристик цільової аудиторії, визначення ключових користувачів і майданчиків, виявлення негативної хвилі обговорень).</p>
3	<p>Лекція 3. Візуальне представлення структур соціальних мереж. Засоби автоматичного розміщення при візуалізації графів</p> <p>1. При аналізі графів соціальних мереж виникає завдання візуального представлення їх структур. Актуальність роботи визначається потребою в програмному комплексі для аналізу і візуалізації графів великих розмірів, що виникає в задачах соціологічного і маркетингового аналізу.</p> <p>2. Архітектура програмного комплексу аналізу графів. Модуль візуалізації графа</p> <p>4. Випадкове розміщення об'єктів. Засоби аналізу графа</p> <p>5. Кругове розміщення об'єктів. Кругове покомпонентне розміщення об'єктів</p> <p>6. Розміщення «одна лінія теми». Розміщення «дві лінії теми»</p> <p>7. Розміщення, засноване на оцінці зв'язності. Розміщення, засноване на виділенні спільнот</p>
4	<p>Лекція 4. Знаходження релевантної інформації. Виведення знань з інформації, яка міститься у INTERNET.</p> <p>Користувачі Інтернет переглядають Веб-сайти або користуються пошуковою сервісами для знаходження необхідної інформації. При використанні пошукових сервісів користувач задає запит за ключовими словами, і відповіддю на цей запит є впорядкований список сторінок, котрі містять задані ключові слова. Недоліки сучасних пошукових засобів: низька релевантність багатьох знайдених ресурсів, незнаходження потрібної інформації через непроіндексованість потрібних сторінок.</p>
5	<p>Лекція 5. Персоналізація інформації. Знання про споживачів / Користувачів.</p> <p>Проблема пов'язана з типом и презентацією інформації, оскільки Користувачі відрізняються тим, яка інформація потрібна їм щоб досягнути своєї мети, яку інформацію шукають и як її краще представити кожному конкретному користувачу.</p>

4.2. План практичних занять

Цілями циклу практичних робіт та контролю засвоєння модулів є закріплення теоретичних знань і набуття практичних навичок в обробці даних соціальних мереж. Структура практичних робіт відображає тенденції у використанні сучасних методів та підходів при аналізі графових та багатомірних БД. Кожна практична робота має *N* варіантів. Номер варіанта студента співпадає із його номером в журналі групи.

№	Тема заняття / план
---	---------------------

№	Тема заняття / план
1-2	<p>Практична робота №1. Інструменти аналізу даних.</p> <p>31. Теоретична підготовка. Аналіз даних на релевантність.</p> <p>32. Захист та перевірка отриманих даних.</p> <p>Мета: під час виконання даної роботи, студент має отримати практичні навички роботи з NetworkX, NumPy, pandas - бібліотеками аналізу даних.</p> <p>NetworkX - це пакет для створення, маніпулювання та вивчення структури, динаміки та функцій складних мереж. Дана бібліотека буде широко використовуватися протягом курсу.</p> <p>NumPy - основний пакет наукових обчислень. Він містить серед іншого: потужний об'єкт N-розмірного масиву; складні функції; інструменти для інтеграції C / C ++ та коду Fortran; корисне лінійну алгебру, перетворення Фур'є та випадкові числа. NumPy є основною бібліотекою для роботи з масивами та матрицями.</p> <p>Pandas - це бібліотека з відкритим кодом, що надає високопродуктивні, прості у використанні структури даних та інструменти аналізу даних для мови програмування Python.</p>
3-4	<p>Практична робота №2. Система обробки слабоструктурованих даних на основі SCRAPY</p> <p>33. Теоретична підготовка. Аналіз даних на релевантність.</p> <p>34. Захист та перевірка отриманих робіт.</p> <p>Мета: під час виконання даної роботи, студент має обраним інструментальним засобом реалізувати програмно інструмент для автоматизованого збору даних з Twitter, реалізувати роботу з базою даних, реалізувати чат-бота для роботи зі створеним інструментом веб-парсингу.</p> <p>Об'єктом для парсингу був обраний Twitter - соціальна мережа, складена з великої кількості мікроблогів. Даний ресурс є зручним для дослідження і вивчення роботи парсерів, тому що формат Twitter передбачає невеликі дописи, розміром не більше 280 символів. Такий формат блогів буде зручним для аналізу твітів (дописів). Для роботи з парсером рекомендовано створити чат-бота на базі месенджера Телеграм, з якого можливо керувати збором даних.</p> <p>Для розв'язання задачі рекомендовано обрати наступні технології: фреймворк Scrapy написаний мовою програмування Python для розробки павука, який збирає дані з веб-сайту, для бази даних було обрано MongoDB, а для створення чат-боту у месенджері Telegram — бібліотеку python-telegram-bot. Проект було створено за допомогою середовища розробки PyCharm, також було використано програмне забезпечення для роботи з базою даних MongoDB Compass.</p>
5-6	<p>Практична робота №3. Статистичні дані мережі</p> <p>35. Теоретична підготовка. Аналіз даних на релевантність.</p> <p>36. Захист та перевірка отриманих робіт.</p> <p>Мета: освоєння практичних навичок роботи з статичними даними соціальних мереж.</p> <p>Направлений мультиграф будується за допомогою класу MultiDiGraph бібліотеки Networkx. Властивості сильної та слабкої зв'язаності графу визначаються за допомогою методів <code>is_strongly_connected</code> та <code>is_weakly_connected</code> відповідно. Посилання №1, посилання №2.</p> <p>Сильнозв'язані та слабкозв'язані компоненти отримуються за допомогою методів <code>strongly_connected_components</code> та <code>weakly_connected_components</code> відповідно. Посилання №1, посилання №2.</p> <p>Найкоротша відстань між вузлами знаходиться за допомогою методу <code>shortest_path_length</code>.</p> <p>Середня відстань між вузлами знаходиться за допомогою методу <code>average_shortest_path_length</code>. Посилання.</p> <p>Ексцентриситетом є найбільша відстань від вершини V до будь-якої іншої вершини. Визначається за допомогою методу <code>eccentricity</code>. Посилання.</p> <p>Діаметр графу є найбільшим ексцентриситетом та визначається за допомогою методу <code>diameter</code>. Посилання.</p> <p>Радіус графу є найменшим ексцентриситетом та визначається за допомогою методу <code>radius</code>. Посилання.</p> <p>Переферією графу є набір вершин, ексцентриситет котрих рівний діаметру графу. Визначається за допомогою методу <code>periphery</code>. Посилання.</p>

№	Тема заняття / план
7-8	<p>Практична робота №4. Міцність з'єднань. Ступінь центральності 37. Теоретична підготовка. Аналіз даних на релевантність. 38. Захист та перевірка отриманих робіт.</p> <p>Мета: навчитися визначати головних учасників мережі та визначати міцність з'єднань мережі. Середній коефіцієнт кластеризації визначається за допомогою методу <code>average_clustering</code>. Посилання.</p> <p>Щоб знайти вершини, які потрібно видалити, для того, щоб роз'єднати граф, використовуйте метод <code>minimum_edge_cut</code>. Посилання.</p> <p>Метод, що перетворює вершини типу <code>String</code> на <code>Int</code> називається <code>convert_node_labels_to_integers</code>. Посилання.</p> <p>Ступінь центральності визначається за допомогою методу <code>degree centrality</code>. Посилання.</p> <p>Центральність за відстанню визначається за допомогою методу <code>closeness centrality</code>.</p>
9-10	<p>Практична робота №5. Міри центральності 39. Теоретична підготовка. Аналіз даних на релевантність. 310. Захист та перевірка отриманих робіт.</p> <p>Мета: дізнатися про міри центральності, їх призначення. Навчитися визначати центральні вузли та на їх основі отримувати головні вузли за допомогою алгоритму Базовий PageRank. В даному курсі лекцій розглядається 3 міри центральності: ступінь зв'язності, ступінь близькості, ступінь посередництва.</p> <p>Ступінь зв'язності базується на кількості з'єднань одного вузла з іншими вузлами в мережі. Бібліотека <code>networkx</code> надає методи для обчислення даної метрики. Для ненаправлених мереж це <code>degree centrality</code> (посилання), а для направлених <code>in_degree centrality</code> (посилання) та <code>out_degree centrality</code> (посилання).</p> <p>Ступінь близькості обчислюється як відношення між кількістю вузлів у мережі - 1 до суми відстаней від обраного вузла до всіх інших у мережі. Базовий PageRank - алгоритм для оцінки важливості вузлів в мережі. В <code>networkx</code> представлений методом <code>pagerank</code> (посилання).</p>

4.3. Завдання для самостійної роботи

Обов'язкові види самостійної роботи

Для опанування матеріалу дисципліни «МЕТОДИ ЗБОРУ ТА ОБРОБКИ ДАНИХ З мережі INTERNET», окрім лекційних та практичних занять – тобто, аудиторної роботи – значну увагу необхідно приділяти самостійній роботі.

Основні види самостійної роботи:

- 1) вивчення додаткової літератури – підготовка інформаційних повідомлень на лекціях;
- 2) виконання індивідуальних завдань.

Теми інформаційних повідомлень співпадають з темами та основними питаннями, які розглядаються на лекціях. В інформаційних повідомленнях також можуть розглядатись новітні засоби та методи в сфері моделювання обчислювальних систем.

Фіксований перелік тем для виконання індивідуальних завдань з дисципліни студентам не пропонується. Теми обираються студентами самостійно та є поглибленням знань про методи та засоби моделювання пошукових систем, які розглядаються в межах дисципліни. Крім того, можуть бути розглянутими деякі специфічні використання методів моделювання, пошуку та аналізу даних в галузях науки, освіти, промисловості, у медицині або спорті тощо. Теми самостійного дослідження узгоджуються з викладачем протягом семестру, до початку залікового тижня.

Приклад завдання для самостійного опрацювання.

Побудувати граф своєї соціальної мережі на базі Facebook

Провести аналіз навантаженості графа.



4.4. Забезпечення освітнього процесу

- Ноутбук, медиа-проектор, екран.
- Комплект слайд-презентацій по курсу.
- Програмне забезпечення для демонстрацій слайд-презентацій PowerPoint.
- Word, Excel
- Wi-Fi

5. Підсумковий контроль

5.1. Типове завдання.

Теорія. 1. Провести аналіз навантаженості графа.

2. Які існують статистичні властивості соціальних мереж?

3. Практичне завдання.

Теоретичні питання 1-2 оцінюються 15 балів, практичне – 3 – в 10 балів.

5.2. Перелік питань підсумкового контролю

1. Для яких цілей використовується аналіз даних з соціальних мереж? Наведіть приклади.
2. Які існують статистичні властивості соціальних мереж?
3. Що таке випадкове блукання? Як воно застосовується для аналізу даних?
4. Що являється спільнотою в соціальних мережах? Назвіть головні методи визначення спільнот в соціальних мережах.
5. Для чого застосовується Марківський алгоритм кластеризації? Наведіть короткий опис алгоритму. Які його плюси та мінуси?
6. Алгоритм Кернігана-Ліна. Короткий опис та застосування.
7. В чому полягає проблема класифікації вузлів графа?
8. Які існують методи класифікації вузлів, засновані на випадковому блуканні?
9. Яким чином виконується збір даних з соціальних мереж для аналізу? Які основні труднощі можуть виникати при зборі даних?

10. Як візуалізація даних допомагає при аналізі даних? Наведіть приклад.
11. Назвіть основні інструменти для візуалізації даних.
12. Які популярні інструменти та бібліотеки можна використовувати для аналізу даних на мові програмування Python.
13. Закони еволюції в соціальних мережах.
14. Що таке структурна візуалізація даних?
15. Що таке статистична візуалізація даних?
16. Як соціальні мережі можуть бути зображені у вигляді зваженого графу?
17. Ітеративний метод класифікації. Короткий опис і застосування.
18. Методи та інструменти для визначення спільнот.
19. Максимізація впливу.
20. Як відстежуються спільноти, що плавно розвиваються?
21. Міри впливу та централізація соціальних мереж.
22. Степеневий розподіл. Його застосування.
23. Дводольний граф. Визначення та застосування.
24. Коефіцієнт кластеризації.
25. Міри відстані.
26. Які існують метрики для вершин графу?
27. Які цілі визначення соціального впливу? Наведіть приклади.
28. Алгоритм поширення міток. Опис та застосування.
29. Статичні та динамічні статистичні властивості соціальних мереж.
30. Як моделюється природа впливу учасників мережі?

6. Критерії оцінювання та засоби діагностики результатів навчання

№	Вид діяльності (завдання)	Максимальна кількість балів
1	Практична робота №1	10
2	Практична робота №2	10
3	Практична робота №3	10
4	Практична робота №4	10
5	Практична робота №5	10
6	Самостійна робота студента	20
	Разом за семестр	70
7	Залік	30
	Всього	100

Критерії оцінювання практичних робіт:

- максимальні **9-10 балів** виставляється за роботу, виконану вчасно та у відповідності до завдання, якщо отримані правильні результати, охайно виконаний звіт, правильно сформульовані висновки до роботи, на захисті продемонстровано розуміння усіх результатів та етапів їх отримання, вільне володіння теоретичним підґрунтям роботи;

- робота оцінюється у **7-8 балів**, якщо маються незначні недоліки при виконанні роботи, отриманих результатах, оформленні звіту, зроблених висновках та при захисті роботи;

- робота оцінюється у **5-6 бали**, якщо маються суттєві недоліки при виконанні роботи, отриманих результатах, оформленні звіту, зроблених висновках та при захисті роботи;

- робота оцінюється у **1-4 бали**, якщо маються значні недоліки при виконанні роботи, отриманих результатах, оформленні звіту, зроблених висновках та при захисті роботи.

Критерії оцінювання самостійної роботи:

- максимальний бал (**18-20 балів**) виставляється за роботу, виконану за відведений час, у відповідності до завдання та у повному обсязі, якщо отримані правильні результати, охайно виконаний звіт, правильно сформульовані висновки до роботи, на захисті продемонстровано розуміння усіх результатів та етапів їх отримання, вільне володіння теоретичним підґрунтям роботи;

- робота оцінюється у **(15-17) бали**, якщо маються незначні недоліки при виконанні роботи, отриманих результатах, оформленні звіту, зроблених висновках та при захисті роботи;
- робота оцінюється у **(11-14) бали**, якщо маються суттєві недоліки при виконанні роботи, отриманих результатах, оформленні звіту, зроблених висновках та при захисті роботи;
- робота оцінюється у **(6-10) бали**, якщо маються значні недоліки при виконанні роботи, отриманих результатах, оформленні звіту, зроблених висновках та при захисті роботи;
- робота оцінюється у **(1-5) бал**, якщо робота виконана самостійно, але не повністю відповідає завданню, маються значні недоліки при виконанні роботи, отриманих результатах, оформленні звіту, або не захищена.

Матеріально-технічне, та/або інформаційне забезпечення

- Проекційне мультимедійне обладнання (проектор, екран, ноутбук/комп'ютер);
- Комп'ютерний клас;
- Доступ до мережі Internet, точка доступу Wi-Fi;
- OS: Windows, Android, iOS;
- Browsers: Chrome / Opera / Mozilla Firefox / MS Edge;
- Програмне забезпечення: Word, Excel, PowerPoint; Skype, Zoom, Google Meet, R, R-Studio; Система електронного навчання Moodle 3.9

7. Рекомендовані джерела інформації

7.1. Основні:

1. О.С. Смирнова, В.В. Шишков. Выбор топологии нейронных сетей и их применение для классификации коротких текстов // International Journal of Open Information Technologies. Том 4, № 8 (2016), с. 50 – 54.
2. Анализ контента социальных медиа в эпоху больших данных // Mebius. 2015-2016. URL: <https://mebius.io/analysis/social-media-content-analysis> (дата обращения: 25.08.2016).
3. Губанов Д.А., Новиков Д.А., Чхартишвили А.Г. Социальные сети: модели информационного влияния, управления и противоборства. – М.: Изд-во физико-математической литературы, 2010. – 228 с.
4. Список систем анализа сетевых структур [Электронный ресурс]. – URL: http://en.wikipedia.org/wiki/Social_network_analysis_software (дата обращения 27.01.2013).

7.2. Додаткові:

1. В.В. Баранюк, А.Д. Десяткова, О.С. Смирнова. Подходы к определению психоэмоциональных особенностей информационного образа пользователя социальных сетей // International Journal of Open Information Technologies. Том 4, № 8 (2016), с. 61 – 65.
2. Барабанов И., Сафронов И., Черненко Е. Разведка ботом // Газета «Коммерсантъ» [Электронный ресурс]. –2012.– №158. – URL: <http://kommersant.ru/doc/2009256> (дата обращения 27.01.2013).
3. Список систем анализа социальных медиа [Электронный ресурс]. – URL: <http://wiki.kenburbary.com> (дата обращения 27.01.2013).
4. Технологии анализа больших данных [Электронный ресурс]. –URL: http://en.wikipedia.org/wiki/Big_data (дата обращения 27.01.2013).
5. Анализ социальных сетей в интернете // Постнаука. 2012-2016. URL: <https://postnauka.ru/longreads/20259#!> (дата обращения: 22.08.2016).

Студентам для вивчення навчального матеріалу надається конспект лекцій з надлишком навчального матеріалу для самостійного опрацювання, а також перелік літератури для засвоєння теоретичного матеріалу.