МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ПРИКАРПАТСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ВАСИЛЯ СТЕФАНИКА

Факультет математики та інформатики

Кафедра алгебри та геометрії

СИЛАБУС НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Інструментальні засоби аналізу та обробки текстів

Освітня програма — Математика комп'ютерних технологій

Спеціальність(ості) — 111 — Математика

Галузь знань — 11 — Математика і статистика

Рівень освіти — Магістр

Затверджено на засіданні кафедри

Протокол №1 від 22.08.2023

Івано-Франківськ — 2023

ЗМІСТ

1. Загальна інформація

2. Контактна інформація

3. Опис дисципліни

4. Структура курсу

5. Система оцінювання курсу

6. Політика курсу

7. Рекомендована література

1. Загальна інформація

Назва дисципліни	Інструментальні засоби аналізу та обробки текстів
Освітня програма	Математика комп'ютерних технологій
Спеціальність	111 — Математика
Галузь знань	11 — Математика і статистика
Освітній рівень	магістр
Статус дисципліни	вибіркова
Рік підготовки / семестр	1-й / 2-й
Обсяг дисципліни	3 кредити
Розподіл за видами занять	Лекції: 14 Практичні: Лабораторні: 16 Самостійна робота: 60
Мова викладання	українська
Посилання на сайт дистанційного навчання	`d-learn.pnu.edu.ua`

2. Контактна інформація

Кафедра	алгебри та геометрії
Викладач(-і)	Гаврилків В.М.
Контактний телефон викладача	59-60-16
E-mail викладача	`volodymyr.gavrylkiv@pnu.edu.ua`
Консультації	Вівторок, 16^{00}

3. Опис курсу

3.1. **Анотація до навчальної дисципліни.** Навчальна дисципліна "Інструментальні засоби аналізу та обробки текстів" займає одне із центральних місць у системі професійної науково-предметної підготовки фахівця з математики комп'ютерних технологій. У цьому курсі вивчаються регулярні мови та їх запис з допомогою регулярних виразів, методи побудови регулярних виразів. Також детально розглянуто застосування регулярних виразів для контекстного пошуку та редагування тексту, використовуючи програму grep, потоковий текстовий редактор sed, текстовий редактор LibreOffice Writer, а також підтримка регулярних виразів у мові Python.

3.2. **Мета і завдання навчальної дисципліни.** Основною метою спецкурсу є ознайомлення студентів з поняттями регулярної мови та регулярного виразу; формування компетентного спеціаліста в даній області, здатного застосовувати і розвивати основні положення дисципліни у науковій і навчальній діяльності, використовувати унікально багаті можливості регулярних виразів у різних прикладних задачах математики та комп'ютерних наук, зокрема при пошуку та обробці текстової інформації.

У результаті вивчення навчальної дисципліни студент повинен

знати:

- поняття формальної мови, регулярної мови та регулярного виразу;
- основні метасимволи, конструкції, квантифікатори;
- стандартні символьні класи;
- поняття групи та посилання;
- основні опції фільтра grep;
- адресацію та функції редактора sed;
- як застосовуються регулярні вирази у LibreOffice Writer;
- як застосовуються регулярні вирази у Python.

вміти:

складати регулярні вирази та використовувати їх для пошуку та обробки тексту в програмах grep, sed, LibreOffice Writer та у мові Python.

4. Структура курсу

Тематика дисципліни						
Назви змістових модулів і тем	Кількість годин					
	вс.	лек.	пр.	лаб.	інд.	сам.
Семестр 2						
Змістовий модуль 1. *Пошук та редагування тексту.*						
Тема 1. *Регулярні мови та регулярні вирази. Розширені регулярні вирази. Застосування регулярних виразів для розширеного контекстного пошуку. Програма grep. Опис опцій. Метасимволи. Символи початку, кінця рядка та довільного символу.* [2, 5, 7, 8, 10]	14	2		2		10
Тема 2. *Вибір. Пошук декількох виразів в одному рядку. Визначення інтервалів та кількості екземплярів. Квантифікатори (повторювачі).* [1, 2, 4, 5, 9]	11	2		2		7
Тема 3. *Символьні класи. Інвертовані символьні класи. Групи та зворотні посилання.* [1, 2, 4, 13]	11	2		2		7
Тема 4. *Застосування регулярних виразів для модифікації тексту. Потоковий текстовий редактор sed. Опис опцій. Адресація. Робочий та допоміжний буфери. Опис функцій редактора sed.*[1, 2, 4, 9]	12	2		2		8
Тема 5. *Функція контекстної заміни. Прапори заміни. Функції видалення, друку та вставки нових рядків. Приклади використання інших функцій.* [2, 4, 9, 13]	11	2		2		7
Тема 6. *Регулярні вирази у програмному продукті LibreOffice Writer.* [1, 2, 4, 13]	13	2		2		9
Тема 7. *Регулярні вирази у мові програмування Python.* [2, 4, 13]	16	2		2		12

Тематика дисципліни						
Назви змістових модулів і тем	Кількість годин					
	вс.	лек.	пр.	лаб.	інд.	сам.
Тема 8. *Контрольна робота.*	2			2		
Всього за модуль:	90	14		16		60
Всього за семестр:	90	14		16		60
Усього годин:	90	14		16		60

5. Система оцінювання дисципліни

Володіння матеріалом дисципліни студенти виявляють при написанні аудиторної контрольної роботи та домашньої контрольної роботи. Максимальна оцінка за кожну з 2-х контрольних робіт становить 40 балів. Домашня контрольна робота передбачає усний захист, включно з формулюваннями використаних означень і фактів.

За активну і змістовну участь у розв'язуванні завдань на лабораторних заняттях здобувач може набрати щонайбільше 20 балів.

Сума балів за семестр визначає підсумкову оцінку згідно поданої нижче таблиці.

Сума балів за всі види навчальної діяльності протягом семестру	Оцінка ЄКТС	Оцінка за національною шкалою
90 – 100	A	відмінно
80 – 89	B	добре
70 – 79	C	добре
60 – 69	D	задовільно
50 – 59	E	достатньо
1 – 49	FX	незадовільно

6. Політика курсу

Самостійне виконання завдань поточного та підсумкового контролю результатів навчання здійснюється під керівництвом викладача, який веде заняття, із наступним їх захистом (для осіб з особливими освітніми потребами ця вимога застосовується з урахуванням їхніх індивідуальних потреб і можливостей). Важливим є надання достовірної інформації про результати власної навчальної (наукової, творчої) діяльності та посилання на джерела інформації у разі використання ідей, відомостей, розробок. Плагіат та інші види академічної недоброчесності не принесуть позитивного результату, а тому не рекомендуються.

Відвідування занять є обов'язковим. Засвоєння пропущеної теми лекції з поважної причини перевіряється під час складання підсумкового контролю. Пропуск лекції з неповажної причини відпрацьовується студентом відповідно вимог кафедри (співбесіда, реферат тощо). Пропущені лабораторні заняття, незалежно від причини пропуску, студент відпрацьовує згідно з графіком консультацій. Поточні "незадовільно", отримані студентом під час засвоєння відповідної теми на лабораторному занятті, перескладаються викладачеві до складання підсумкового контролю з обов'язковою відміткою у журналі обліку роботи академічних груп.

7. Рекомендована література

1. Бєлов Ю.А. Інструментальні засоби програмування: навчальний посібник / Ю.А. Бєлов, В.С. Проценко, П.Й. Чаленко. – К.: Либідь, 1993. – 248 с.
2. Гаврилків В.М. Регулярні вирази у програмних продуктах: навчальний посібник / В.М. Гаврилків. – Івано-Франківськ: Голінеї О.М., 2012. – 72 с.

3. Гаврилків В.М. Формальні мови та алгоритмічні моделі: навчальний посібник / В.М. Гаврилків. – Івано-Франківськ: Голіней О.М., 2012. – 172 с.
4. Гаврилків В.М. Програмні засоби розбору та обробки тексту / Івано-Франківськ: Голіней, 2013. - 41 с.
5. Гаврилків В.М. Формальні мови та алгоритмічні моделі: навчальний посібник (вид. друге) / В.М. Гаврилків. – Івано-Франківськ, 2023. – 180 с.
6. Гаврилків В.М. Практичні методи розробки компіляторів / В.М. Гаврилків. – Івано-Франківськ, 2023. – 71 с.
7. Матвієнко М.П. Дискретна математика / М.П. Матвієнко. – К.: Ліра-К, 2013. – 324 с.
8. Трохимчук Р.М. Дискретна математика: навч. пос. / Р.М. Трохимчук. – К.: "Персонал 2010. – 528 с.
9. Bambenek J. grep Pocket Reference / John Bambenek and Agnieszka Klus. – O'Reilly Media, 2009 – 75 p.
10. Ding-Zhu Du. Problem Solving in Automata, Languages, and Complexity / Ding-Zhu Du, Ker-I Ko. – New York: WIP, 2001. – 388 p.

Додаткова література

11. Бондаренко М.Ф. Комп'ютерна дискретна математика: підручник / М.Ф. Бондаренко, Н.В. Білоус, А.Г. Руткас. – Харків: «Компанія СМІТ», 2004. – 480 с.
12. Нікольський Ю.В. Дискретна математика / Ю.В. Нікольський, В.В. Пасічник, Ю.М. Щербина. – К.: Видавнича група BHV, 2007. – 368 с.
13. Goyvaerts J. Regular Expressions Cookbook / Jan Goyvaerts and Steven Levithan. – O'Reilly Media, 2009 – 494 p.

Викладач Гаврилків В.М.