

Міністерство освіти і науки України  
Львівський національний університет імені Івана Франка  
Кафедра оптоелектроніки та інформаційних технологій

**«ЗАТВЕРДЖУЮ»**

Декан факультету електроніки та  
комп'ютерних технологій

\_\_\_\_\_ Фургала Ю. М.

« \_\_\_\_\_ » \_\_\_\_\_ 2020 року

## **КОМП'ЮТЕРНА ЛІНГВІСТИКА**

### **ПРОГРАМА**

**навчальної дисципліни**

**підготовки бакалаврів**

**галузі знань 12 Інформаційні технології**

**спеціальності 126 Інформаційні системи та технології**

**факультету електроніки та комп'ютерних технологій**

**(шифр за ОП: НД 1.1.02)**

2020 рік

РОЗРОБЛЕНО ТА ВНЕСЕНО:

Львівським національним університетом імені Івана Франка

РОЗРОБНИКИ ПРОГРАМИ:

Кушнір О. С., докт. фіз.-мат. наук, проф., зав. кафедри оптоелектроніки та інформаційних технологій факультету електроніки та комп'ютерних технологій

Затверджено на засіданні Вченої ради факультету електроніки та комп'ютерних технологій

«\_\_» \_\_серпня\_\_ 2020 року, протокол №\_7/10\_\_

Голова Вченої ради

Фургала Ю. М.

Обговорено та рекомендовано до затвердження Навчально-методичною радою факультету електроніки та комп'ютерних технологій

«\_\_» \_\_серпня\_\_ 2020 року, протокол №\_4\_

Голова Навчально-методичної ради  
факультету електроніки та комп'ютерних технологій

Лучечко А. П.

Програму схвалено на засіданні кафедри оптоелектроніки та інформаційних технологій

«\_\_\_\_\_» \_\_серпня\_\_ 2020 року, протокол № \_\_\_\_\_

Завідувач кафедри оптоелектроніки  
та інформаційних технологій

Кушнір О. С.

© Кушнір О. С., 2020 рік

## ЗАГАЛЬНІ ВІДОМОСТІ

Програма вивчення навчальної дисципліни «Комп'ютерна лінгвістика» складена відповідно до освітньо-професійної програми підготовки здобувачів першого (бакалаврського) ступеня вищої освіти за спеціальністю **126 Інформаційні системи та технології** і є нормативною навчальною дисципліною.

**Міждисциплінарні зв'язки.** Для вивчення цієї дисципліни потрібні знання, одержані за результатами вивчення загальних і спеціальних дисциплін відповідних галузі знань і спеціальності, насамперед «Вища математика», «Дискретна математика», «Теорія ймовірності та математична статистика», «Програмування», «Бази даних» та ін.

**Предметом** вивчення навчальної дисципліни є основи комп'ютерної та статистичної лінгвістики.

Програма дисципліни складається з одного модуля та трьох змістових модулів.

### 1. МЕТА ТА ЗАВДАННЯ НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

**1.1. Метою викладання навчальної дисципліни** є одержання студентами необхідних теоретичних знань з основ комп'ютерної лінгвістики, а також формування в студентів практичних навичок, які б дали змогу ефективно застосовувати засвоєні знання та методи комп'ютерної лінгвістики.

**1.2. Основні завдання вивчення дисципліни:** ознайомити студентів з основними поняттями лінгвістики, ідеями, підходами та методами комп'ютерної лінгвістики, структурним підходом у комп'ютерній лінгвістиці, класифікацією та особливостями продуктів комп'ютерної лінгвістики, проблемами автоматичного аналізу та розпізнавання мови, синтезом мови та мовними технологіями, основами машинного перекладу та комп'ютерної лексикографії.

**1.3. Згідно з вимогами освітньо-професійної програми студенти повинні знати:**

- поняття, методи та продукти комп'ютерної лінгвістики
- основні методи комп'ютерної лінгвістики;
- теоретично аналізувати основні теорії та моделі комп'ютерної лінгвістики

**вміти:**

- класифікувати та вирішувати типові задачі комп'ютерної лінгвістики;
- теоретично аналізувати основні теорії та моделі комп'ютерної лінгвістики
- працювати з відповідними програмними продуктами;
- застосовувати комп'ютерну техніку для вирішення лінгвістичних задач;
- виробляти та реалізувати лінгвістичні комп'ютерні алгоритми;
- користуватися програмами для машинного перекладу.

Форма навчання	Семестр	Всього кредитів/годин	Розподіл навчального часу за видами занять <sup>1</sup>					Семестрова атестація
			Лекції	Практичні заняття	Семінарські заняття	Лабораторні роботи	С.Р.С.	
Денна	7	4/120	32	–	–	32	56	іспит

## II. ІНФОРМАЦІЙНИЙ ОБСЯГ НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

*Змістовий модуль 1. Поняття, методи та продукти комп'ютерної лінгвістики*

**Тема 1.1.** Вступ. Лінгвістика та її структура.

**Тема 1.2.** Основні поняття лінгвістики.

**Тема 1.3.** Розвиток ідей, теорій, підходів і методів комп'ютерної лінгвістики.

**Тема 1.4.** Продукти комп'ютерної лінгвістики: досягнення та проблеми.

*Змістовий модуль 2. Основні теорії та моделі, покладені в основу алгоритмів комп'ютерної лінгвістики*

**Тема 2.1.** Мова як двонаправлений перетворювач змісту та тексту.

**Тема 2.2.** Лінгвістичні знаки.

**Тема 2.3.** Лінгвістичні моделі.

*Змістовий модуль 3. Аналіз, розпізнавання та синтез природної мови, машинний переклад та комп'ютерна лексикографія*

**Тема 3.1.** Автоматичне введення звуків мови, аналіз та розпізнавання мови комп'ютером.

**Тема 3.2.** Синтез мови та мовні технології.

**Тема 3.3.** Машинний переклад і комп'ютерна лексикографія.

## III. ПРИБЛИЗНИЙ ПЕРЕЛІК ЛАБОРАТОРНИХ РОБІТ

1. Розрахунок статистичних параметрів текстів
2. Програмування задачі визначення співвідношення кількості голосних і приголосних (прозорості мови) для українських текстів мовою C
3. Програмування розпізнавання лінгвістичних одиниць (літер і знаків, слів, частин мови).
4. Вивчення роботи програмного пакету FineReader.
5. Програмування задачі відновлення вихідних словоформ іменників.
6. Програмування простих задач морфологічного синтезу дієслівних форм.
7. Побудова дерев синтаксичних залежностей для стверджувальних речень.
8. Генерування фрагментів тексту: формальні контекстно-вільні граматики.
9. Робота з базами лінгвістичних даних. Побудова запитів з урахуванням лінгвістичних міркувань.
10. Обробка природних мов у рамках середовища Python.
11. Вивчення роботи електронного словника Lingvo.
12. Програмний пакет Stylus: налаштування інтерфейсу та робота.

## IV. РЕКОМЕНДОВАНА ЛІТЕРАТУРА

### Базова

1. Bolshakov I. Computational linguistics. Models, resources, applications / I. Bolshakov, A. Gelbukh. – Mexico : Ciencia de la Computacion, 2004. – 198 p.
2. Волошин В. Г. Комп'ютерна лінгвістика / В. Г. Волошин. – Суми : Університетська книга, 2004. – 382 с.
3. Шемакин Ю. И. Начала компьютерной лингвистики / Ю. И. Шемакин. – Москва : Изд-во МГОУ, 1992. – 182 с.
4. Баранов А. Н. Введение в прикладную лингвистику / А. Н. Баранов. – Москва : Эди-ториал УРСС. – 2001. – 360 с.
5. Мирам Г. Алгоритмы перевода: Вступительный курс по формализации перевода / Г. Мирам. – Киев : Эльга, Ника-Центр, 2004. – 176 с.

6. Хархалис Р. И. Компьютерный перевод иностранных текстов / Р. И. Хархалис. Киев : Терези, 1998. – 193 с.
7. Белоногов Г. Г. Компьютерная лингвистика и перспективные информационные технологии / Г. Г. Белоногов. – Москва : Русский мир, 2004. – 248 с.
8. Большакова Е. И. Автоматическая обработка текстов на естественном языке и компьютерная лингвистика / Е. И. Большакова, Э. С. Клышинский, Д. В. Ландэ, А. А. Носков, О. В. Пескова, Е. В. Ягунова. – Москва : МИЭМ, 2011. – 272 с.

#### **Допоміжна**

1. Пиотровский Р. Г. Математическая лингвистика / Р. Г. Пиотровский, К. Б. Бектаев, А. А. Пиотровский. – Москва : Высшая школа, 1977. – 383 с.
2. Заде Л. А. Понятие лингвистической переменной и его применение к принятию приближенных решений / Л. А. Заде. – Москва : Мир, 1976. – 165 с.
3. Фланаган Дж. Л. Анализ, синтез и восприятие речи / Дж. Л. Фланаган. – Москва : Связь, 1968. – 398 с.
4. Маркел Дж. Д. Линейное предсказание речи / Дж. Д. Маркел, А. Х. Грэй. – Москва : Связь, 1980. – 308 с.
5. Рассел С. Искусственный интеллект. Современный подход / С. Рассел, П. Норвинг. – Москва : Изд. дом «Вильямс», 2006. – 1408 с.
6. Анисимов А. В. Компьютерная лингвистика для всех: Мифы. Алгоритмы. Язык / А. В. Анисимов. – Киев : Наукова думка, 1991. – 208 с.

#### **V. ФОРМА ПІДСУМКОВОГО КОНТРОЛЮ УСПІШНОСТІ НАВЧАННЯ**

Контроль засвоєння матеріалу включає поточний контроль (дві модульні контрольні роботи  $2 \times 10 = 20$  балів), лабораторні роботи (40 балів) і оцінювання знання теоретичного матеріалу на іспиті (40 балів). Сумарна оцінка виставляється за 100-бальною шкалою.

Вивчення курсу завершується іспитом, який проводиться в письмовій формі.

#### **VI. ЗАСОБИ ДІАГНОСТИКИ УСПІШНОСТІ НАВЧАННЯ**

Для засвоєння теоретичного матеріалу, підготовки до виконання лабораторних завдань студентам надається можливість користуватися бібліотеками Львівського національного університету імені Івана Франка, а студентам старших курсів (починаючи з третього) – Львівською національною науковою бібліотекою України імені В. Стефаника. Студенти мають змогу отримати консультації з питань дисципліни в лектора та викладачів, які проводять лабораторні заняття.

Оцінка якості засвоєння навчальної дисципліни включає поточний контроль успішності та складання підсумкового екзамену.

Для поточного контролю засвоєння студентами навчального матеріалу передбачається виконання та захист 12 обов'язкових лабораторних робіт і написання 2 модульних контрольних робіт.

При оцінюванні лабораторної роботи враховується підготовка до виконання лабораторної роботи, хід виконання лабораторної роботи, оформлення звіту, отримані результати та захист звіту про виконану лабораторну роботу.