**Перелік питань**

з навчальної дисципліни «Комп’ютерний аналіз текстових даних»

за спеціальністю 035 Філологія

освітнього ступеню «бакалавр»

|  |  |
| --- | --- |
| № з/п | Зміст питання |
|  | Загальні поняття Python. |
|  | Робота зі списками, кортежами, рядками, словниками |
|  | Умовний оператор. |
|  | Цикли for, while. |
|  | Підключення модулів. |
|  | Функції. |
|  | Файли. Запис та зчитування інформації у файл. |
|  | Побудова графіків та діаграм в Matplotlib. |
|  | Базові поняття штучного інтелекту в лінгвістиці |
|  | Text Mining як основа аналізу неструктурованої текстової інформації |
|  | Регулярні вирази. |
|  | Основи NLP для тексту. |
|  | Токенізація за реченнями та словами. |
|  | Видалення стоп-слів. |
|  | Нормалізація слів. |
|  | Стемінг. |
|  | Лематизація. |
|  | Використання нормалізованих даних для аналізу тексту. |
|  | Векторизація текстів. |
|  | Мішок слів. TF-IDF. |
|  | N – грами |
|  | Колокації |
|  | Морфологічний аналіз. |
|  | Аналіз тональності. |
|  | Робота з корпусами засобами бібліотеки NLTK |
|  | Аналіз корпусів. |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |