**Київський національний університет імені Тараса Шевченка**

**факультет радіофізики, електроніки та комп’ютерних систем**

лабораторна робота № 1

**Тема:** «Дослідження кількості інформації при різних варіантах кодування»

Роботу виконав

студент ІІІ курсу

КІ-МА

Грищук Олександр

Київ 2020**Хід виконання роботи:**

1. Дослідження кількості інформації в тексті
   1. Оберіть 3 текстових файла різного тематичного та лінгвістичного спрямування

Sample1.txt – вірш Т. Г. Шевченка «Думи мої думи»

Sample2.txt – фрагмент статті про PHP з lurkore.to

Sample3.txt – RFC 2795 (IMPS)

* 1. Створіть програму (будь-якою зручною для вас мовою), яка в якості вхідних даних приймає текстовий файл, та аналізуючи його вміст:
     1. обраховує частоти (імовірності) появи символів в тексті
     2. обраховує середню ентропію алфавіту для даного тексту
     3. виходячи з ентропії визначає кількість інформації та порівнює її з розмірами файлів
     4. виводить на екран значення частот, ентропії та кількості інформації 4. Проведіть стиснення кожного вхідного файлу за допомогою 5 різних алгоритмів стиснення (zip, rar, gzip, bzip2, xz, або будь-які інші на ваш вибір, можна використовувати готові програмні засоби для стиснення).
  2. Порівняйте результуючі обсяги архівів з обчисленою кількістю інформації та ​ наведіть у звіті висновки ​ щодо кореляції цих величин для обраних вами файлів (яка відмінність, що вийшло більше і чому)

1. <Запитання>

<Відповідь>

….

**Висновки:** <висновки до лабораторної роботи>